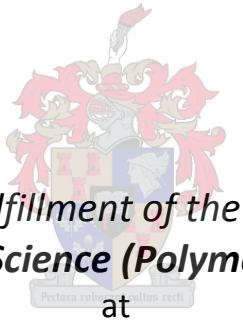
Multidimensional Fractionation of Wood-Based Tannins

By

Nonhlanhla Mtandi Radebe



Thesis presented in partial fulfillment of the requirements for the degree
Master of Science (Polymer Science)
at
University of Stellenbosch

Supervisor: Prof. Harald Pasch
Faculty of Science
Department of Chemistry and Polymer Science

March 2011

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

March 2011

Abstract

High molar mass tannin extracts are complex mixtures which are distributed in both molar mass and chemical composition. Condensed tannins from quebracho and mimosa and hydrolysable tannins of tara, chestnut wood and turkey gall were studied. Application of a single analytical technique is not sufficient to elucidate the complete structures present in the extracts. ^{13}C Nuclear Magnetic Resonance (NMR) spectroscopy and Matrix Assisted Laser Desorption/Ionisation Time-of-Flight (MALDI-TOF) mass spectrometry were applied in order to determine the chemical composition and molar mass, respectively. A new mass spectrometric method that can uniquely determine the oligomer microstructure was developed using Collision Induced Dissociation (CID) experiments. Bulk analysis only showed the average composition of the extracts, in order to obtain specific information on the molar mass and chemical composition distributions. Hydrophilic Interaction Liquid Chromatography (HILIC) was used for analysis of the condensed tannins and for the hydrolysable tannins Normal Phase Liquid Chromatography (NP-LC) was utilised. The HILIC separation was up-scaled and the fractions were collected and analysed by MALDI-TOF, and this coupling revealed that separation occurs by molar and chemical composition. For separation of the molecules only by size, Size Exclusion Chromatography (SEC) analyses were carried out; this allowed for relative comparison of the tannin molecules. In conclusion, for characterisation of high molar mass tannins a multi-dimensional approach was necessary since the various distributions present in these extracts are superimposed.

Opsomming

Hoë molekulêre massa tannienekstrakte is komplekse mengsels, in terme van beide molekulêre massa en chemiese samestelling. Gekondenseerde tanniene vanaf quebracho en mimosa, en hidroliseerbare tanniene vanaf tara, kastaaiinghout en Turksegal is bestudeer. Die gebruik van ‘n enkele analitiese tegniek is nie voldoende om die volledige struktuur van komponente teenwoordig in die ekstrakte te analiseer nie. ^{13}C KMR-spektroskopie en MALDI-TOF-massaspektroskopie is gebruik om die chemiese samestelling en molekulêre massa, onderskeidelik, te bepaal.

‘n Nuwe metode is ontwikkel vir die bepaling van die oligomeer-mikrostruktuur deur gebruik te maak van botsings-geïnduseerde dissosiasie eksperimente. Grootmaat analise het net die gemiddelde samestelling van die ekstrak bepaal. Hidrofiliese-interaksie-vloeistofchromatografie (HILIC) is gebruik vir die analise van gekondenseerde tanniene en gewone fase-vloeistofchromatografie is gebruik vir die hidroliseerbare tanniene. Die HILIC-skeiding is op groter skaal uitgevoer en die fraksies is versamel en gebruik vir MALDI-TOF analise. Hierdie koppeling het getoon dat skeiding plaasvind op grond van molekulêre massa en chemiese samestelling.

Grootte-uitsluitingschromatografie is gebruik vir die skeiding van molekules alleenlik op grootte. Hierdeur kon ‘n relatiewe vergelyking van die tannienmolekules gemaak word.

Vir die karakterisering van hoë molekulêre massa tanniene is ‘n multi-dimensionele benadering nodig aangesien die verskeie verspreidings teenwoording in hierdie ekstrakte supergeponeerd is.

Acknowledgements

I would like to express my sincerest gratitude to all the people and institutions that supported me throughout the two years of my MSc. In particular I want to thank Prof. Harald Pasch for his continued support, contributions and funding from the beginning of this project, and from him I have learnt a lot.

I would also like to thank Dr. James Mcleary for offering up his time to secure financial support for this project. I am very much thankful for the funding that Plascon and National Research Foundation provided me in order to perform this research. I also want to thank Deutsches Kunststoff-Institut (DKI) in Germany for giving me the opportunity to use their MALDI-TOF instrument, and especially the group members who made my stay a very enjoyable one, thank you.

I have to express my heartfelt gratitude to my family and friends. Although your contribution was not a direct one, I would like to thank you for all your support and guidance for the duration of my studies. It would not suffice not to thank my group members, new and old; thank you for all your assistance and advice when I needed it; it has been a real honour knowing you. I want to specifically acknowledge Nadine Pretorius who has offered continued support and mentorship; the long conversations and lab sessions offered for this MSc and honours projects allowed me to attain the skills that I have now acquired.

And last, but not least I would like to acknowledge and honour the Lord God Almighty without whom any of this would not have been possible.

Table of Contents

List of Abbreviations and Symbols

ACN	Acetonitrile
AcOH	Acetic Acid
APCI	Atmospheric Pressure Chemical Ionisation Interface
CCD	Chemical Composition Distribution
CID	Collision Induced Dissociation
D_2O	Deuterated Water
DMAc	Dimethyl Acetamide
DMF	Dimethyl Formamide
DP	Degree of Polymerisation
ELSD	Evaporative Light Scattering Detector
ESI-MS	Electrospray Ionisation Mass Spectrometry
FAB	Fast Atom Bombardment Mass Spectrometry
FTIR	Fourier Transform Infrared Detector
HHDP	Hexahydroxydiphenic Acid
HILIC	Hydrophilic Interaction Liquid Chromatography
HPLC	High Performance Liquid Chromatography
K_{ads}	Distribution Coefficient of Adsorption
K_D	Chromatographic Distribution Coefficient
K_{sec}	Distribution Coefficient of Steric Exclusion
LAC	Liquid Adsorption Chromatography
LC	Liquid Chromatography
LiCl	Lithium Chloride
MALDI-TOF	Matrix Assisted Laser Desorption/Ionisation Time-of-Flight mass spectrometry
MeOH	Methanol
MMD	Molar Mass Distribution
Mn	Number average Molar mass
Mw	Weight average Molar mass
$NaHSO_3$	Sodium Metabisulphite
NaOH	Sodium Hydroxide
NMR	Nuclear Magnetic Resonance Spectroscopy
NP-LC	Normal Phase Liquid Chromatography
PC	Procyanidins
PD	Prodelphinidins
PDA	Photodiode Array Detector
PDI	Polydispersity Index
PEI	Polyethyleneimine
PGG	Pentagalloylglucose
PMMA	Polymethyl Methacrylate
PS	Polystyrene
PSD	Post Source Decay
Q-TOF	Quadrupole Time-of-Flight
RDA	Retro-Diels-Alder Fission
RI	Refractive Index Detector
RP-LC	Reversed Phase Liquid Chromatography

SEC	Size Exclusion Chromatography
SPE	Solid Phase Extraction
T	Tempreature
THF	Tetrahydrofuran
UV	Ultraviolet Detector
V_a	Volume of Adsorption Layer
V_i	Interstitial Volume
V_p	Pore Volume
V_r	Retention Volume
V_s	Volume of Stationary Phase
V_{stat}	Total ‘Stationary’ Volume
WET	Water suppression through T1 effects (NMR solvent suppression technique)
ΔG	Change in Gibbs Free Energy
ΔH	Change in Enthalpy
ΔS	Change in Entropy
Φ	Eluent Composition

Chapter 1

General Introduction

Tannins are defined as water-soluble polyphenolic molecules with defined features and are obtained from various plants extracts [1]. These molecules are widely distributed in nature and may be located in the stem, leaves or fruit of a plant. They are divided into two distinct groups, namely, hydrolysable and condensed tannins. The interest in these materials arises from their wide availability but limited use since their exact structure is still not understood. The physiological activity of tannins has been studied extensively for decades and their industrial importance less so. The health benefits of tannin molecules contained in food products include anti-oxidant, anti-inflammatory and anti-microbial properties among others [1-3]. Commercially extracted tannins are used to synthesise wood adhesives, super-plasticizers and are used in leather tanning [4]. Wood plants are known to be composed of high molar mass tannins and these are the molecules that are extracted on a commercial scale [4-6]. The molecules are extracted by different procedures and may be modified to synthesise intermediates suitable for further reactions [7,8]. The tannins can thus be distributed both in molar mass and chemical composition. The function of the tannins has been shown to be affected by the degree of polymerisation and structure to some extent [9]. In the case of the condensed tannins variation of chemical composition is a result of the distribution of the –OH groups on the phenolic moieties. **Figure 1.1** shows the possible structures that may be contained in such extracts. In the case of hydrolysable tannins the chemical composition difference is a result of the extent of oxidation and substitution on the carbohydrate core of the molecule, the generalised structure is indicated in **Figure 1.1**.

The structure-property relationship of the tannins is significant since the tannin molecules are contained in such a vast number of plant species. Understanding this relationship may allow greater use of this natural resource. Tannins with the same chemical composition and molar mass perform the same function and if the structure can be understood similarities in composition between extracts obtained from different plant sources may be discovered. Unlike in the case of synthetic materials the composition of the tannin extract cannot be controlled, and the pursuit can only be to understand these complex samples.

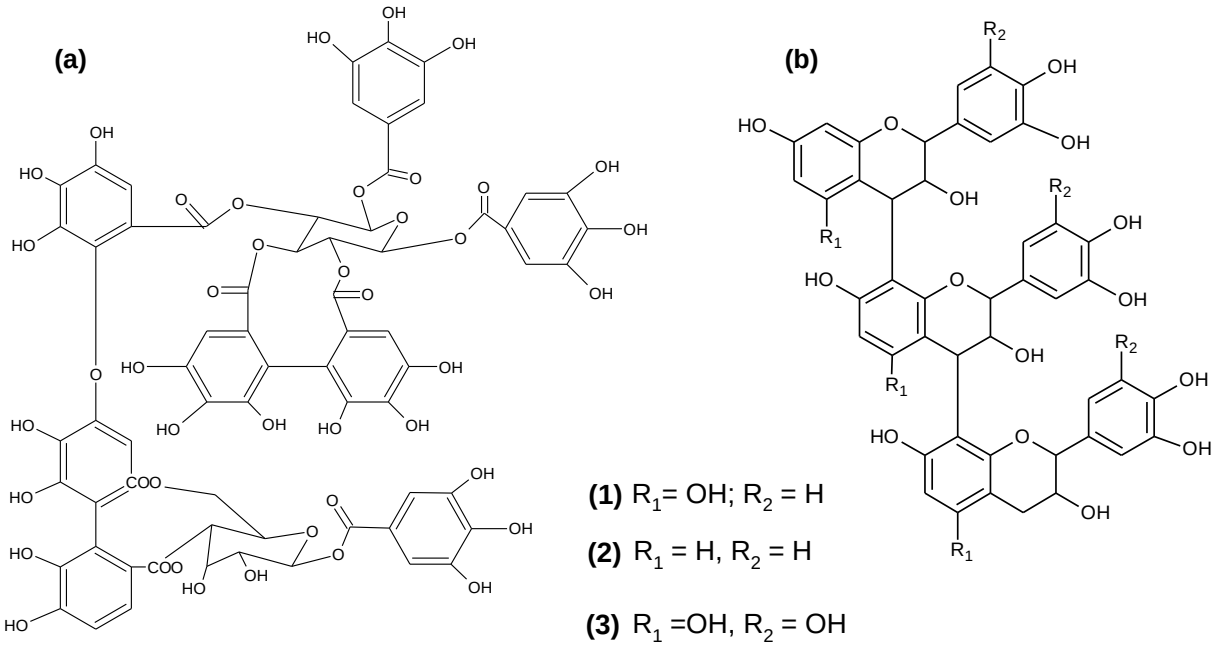


Figure 1.1: Generalised structures of (a) hydrolysable tannins (b) condensed tannins. Variations of these molecules may be present: (1) Catechin (procyanidins), (2) fisetinidin (profisetinidins), and (3) gallocatechin (prodelphinidins) [5,10].

In order to understand the described relationship, full characterisation is required, that includes not only the averaging techniques but methods that are able to distinguish between the various distributions present. The most powerful method used for this purpose has been High Performance Liquid Chromatography (HPLC) for tannins extracted from consumable products. This method has been used in the Reversed Phase (RP) mode to determine the chemical composition and Normal Phase (NP) modes are applied to separate the molecules according to their degree of polymerisation [9,11-13]. In polymer analysis the well developed method for separating molecules by size is Size Exclusion Chromatography (SEC) and this has been applied to analyse methylated and acetylated derivatives of tannin molecules [14]. A relatively new NP method that has gained popularity in the separation of oligomeric condensed tannins by degree of polymerisation is Hydrophilic Interaction Liquid Chromatography (HILIC) which makes use of a diol stationary phase [15,16].

A powerful but relatively simple technique that is able to determine structural properties of oligomeric tannins is Matrix-Assisted Laser Ionisation/Desorption Time-of-Flight (MALDI-TOF) mass spectrometry [7,8,17].

MALDI-TOF analysis has been applied extensively in the analysis of specifically condensed tannins. Tannins contain high frequency of isomeric structures and thus unambiguous assignment of the peaks to specific oligomers has not been possible, the structures elucidated are based on information obtained from other techniques. A relatively new technique that is available to determine the monomer sequence of oligomers is MALDI-TOF equipped with a collision cell and the technique is referred to as MALDI-TOF Collision Induced Dissociation (CID) mass spectrometry. This technique can be used in order to determine the monomer sequence in an oligomer chain by fragmenting the molecule into its constituents [18]. Therefore it is capable of giving the direct microstructure of each oligomer as detected in MALDI-TOF. This technique is used mainly in analysis of oligosaccharides in order to determine the sequence of the monomer units, application to tannins to the best of our knowledge has not been attempted.

Although the MALDI-TOF technique on its own provides only mass information, in combination with liquid chromatographic separations it can provide oligomer specific information. MALDI-TOF has been combined with LC separations in order to determine the chemical composition, however, in the case of tannin analysis this combination has been limited to fractions obtained from incomplete preparative separations [19]. These preparative methods are applied as a clean-up step prior to further analysis and thus the separation is incomplete, that is, the molecules are often separated into three main fractions, the monomeric, oligomeric and polymeric fractions.

As described earlier, tannin extracts are complex and thus a simple analysis by a single method is insufficient to fully describe the constituents of each tannin extract. In polymer analysis combination of MALDI-TOF analysis with an SEC separation can give further insight into the content of the fraction analysed [20]. A preparative HILIC method was shown by Kelm et al. to be able to separate the cacao tannin oligomers by size, detection in this case was carried out by off-line ESI-MS [15]. The fractions were shown to be composed of distinct oligomers and overlap was minimal. This method was able to determine the chemical compositions contained in the cacao tannin extract.

The aim of this thesis was to develop comprehensive analytical methods able to determine the chemical composition and molar mass distributions present in tannin extracts. The focus was on commercial wood

tannin extracts which are known to be composed of higher molar masses than those usually present in the consumable products. Analytical methods available for characterisation these types of tannins are limited. This document is divided into two main sections, the first focussing on the characterisation of condensed tannins and in the second part; the analysis of hydrolysable tannins is described.

In the first experimental chapter, spectroscopic techniques used to compare the condensed tannin extracts are shown. The condensed tannins from quebracho and mimosa woods obtained by various extraction methods are analysed. Bisulphited samples from quebracho have been shown to be subject to hydrolysis and sodium hydroxide/maleic anhydride modification can also replace some of the –OH groups on the phenolic moieties with methyl groups [7,21]. Solvent extraction in tannins removes the polymeric fraction which may be present [7]. Two mimosa tannin samples were considered: the first was water-extracted and the second was extracted with bisulphited water. The point was to compare commercial extracts obtained by the various methods first with each other and then with other extracts. The cacao tannin was used as a reference since it has a known chemical composition. Spectroscopic techniques such as ^{13}C NMR, MALDI-TOF and ESI-MS should enable determination of any differences present in the extracts. Further structural determinations were carried out with MALDI-TOF CID which allowed monomer sequence determination. HILIC separations are known to provide invaluable information and thus the separations obtained for the quebracho and cacao tannin were coupled to MALDI-TOF.

The second experimental section was dedicated to the analysis of hydrolysable tannins. The commercial gallotannins, tara and turkey gall were considered as well as an ellagitannin (chestnut tannin). All the samples were solvent extracted. The structures present in all of the three extracts were compared to information available in literature and thus the difference in structure was related to the mode of extraction. The same types of analyses were carried out for these tannins as was done for the condensed tannins. The HPLC separations were carried out in normal phase mode.

References

[1] E. Haslam, Practical polyphenolics from structure to molecular recognition and physiological action, Cambridge University Press, New York, 1998.

[2] L.L. Zhang, Y.M. Lin, Molecules 13 (2008) 2986.

[3] M. Noferi, E. Masson, A. Merlin, A. Pizzi, X. Deglise, J. Appl. Polym. Sci. 63 (1997) 475.

[4] A. Pizzi, in M. Belgacem, A. Gandini (Editors), Monomers, Polymers and Composites from Renewable Resources, Elsevier, 2008, p. 179.

[5] A. Pizzi, Wood Adhesives Chemistry and Technology, Marcel Dekker, New York, 1983.

[6] C.S. Ku, S.P. Mun, Wood Sci. Technol. 41 (2007) 235.

[7] H. Pasch, A. Pizzi, K. Rode, Polymer 42 (2001) 7531.

[8] A. Pizzi, H. Pasch, K. Rode, S. Giovando, J. Appl. Polym. Sci. 113 (2009) 3847.

[9] V. Cheynier, J.-M. Souquet, E. Le Roux, S. Guyot, J. Rigaud, Methods Enzymol. 299 (1999) 178.

[10] A. Pizzi, Advanced Wood Adhesives Technology CRC Press, New York, 1994.

[11] J.F. Hammerstone, S.A. Lazarus, A.E. Mitchell, R. Rucker, H.H. Schmitz, J. Agric. Food. Chem. 47 (1999) 490.

[12] B. Zywicki, T. Reemtsma, M. Jekel, J. Chromatogr. A 970 (2002) 191.

[13] T. Okuda, T. Yoshida, T. Hatano, J. Nat. Prod. 52 (1989) 1.

[14] A. Yanagida, T. Shoji, T. Kanda, Biosci. Biotechnol., Biochem. 66 (2002) 1972.

[15] M.A. Kelm, J.C. Johnson, R.J. Robbins, J.F. Hammerstone, H.H. Schmitz, J. Agric. Food. Chem. 54 (2006) 1571.

[16] K.M. Kalili, A. de Villiers, J. Chromatogr. A 1216 (2009) 6274.

[17] P. Navarrete, A. Pizzi, H. Pasch, K. Rode, L. Delmotte, Ind. Crops Prod. 32 (2010) 105.

[18] A.I.T. Jackson, K.R. Jennings, J.H. Scrivens, J. Am. Soc. Mass. Spectrom. 8 (1997) 76.

[19] C. Perret, R. Pezet, R. Tabacchi, Phytochem. Anal. 14 (2003) 202.

[20] J.-A. Raust, A. Brull, C. Moire, C. Farcet, H. Pasch, J. Chromatogr. A 1203 (2008) 207.

[21] A. Pizzi, D. Thompson, J. Appl. Polym. Sci. 55 (1995) 107.

Chapter 2

Theoretical Considerations

2.1. Introduction

The definition of what constitutes a tannin has varied greatly over the years; a 'tannin' used to define any plant extract that had the ability to tan leather. Historically this was the only application of these plant extracts. The tannins are contained in high concentrations in different plant species. In general they can be obtained from all parts of the plant, i.e. the stem, leaves, and fruit of the plant. Recently, however, extensive research has been carried out on the identification, application and structure-property relationships of various plant extracts and the tannins have been clearly defined as water-soluble polyphenolic molecules with certain features obtained from plant extracts. They are further divided into two distinct groups, depending on their structure. The first is referred to as polyflavonoids or proanthocyanidins and the second group is named hydrolysable tannins.

The polyflavonoids obtain their name from the fact that they are formed from flavanol units bonded at the C6-C3-C6 or C6-C4-C8 positions and the proanthocyanidin name is based on the fact that in acidic medium these molecules break up to form anthocyanidins which are red coloured compounds. The hydrolysable tannins on the other hand obtained their name from the fact that they hydrolyse in acidic or basic medium into sugars, simple phenols and their carboxylic acids [1].

A more extensive review of the history and application of tannins is covered by Pizzi [2]. Tannins went through a phase whereby they were exclusively employed in leather tanning, however, in the last decades tannins were replaced by synthetic materials. It followed then that other applications would need to be found for these plant extracts and thus they were utilised to synthesise wood adhesives, later used as flocculants and subsequently used as super-plasticisers [2,3]. Tannin containing compounds were used in folk medicine before the active ingredients were discovered. In recent years extensive studies have been done on these medicines and their benefits have been defined [4,5]. The function of tannins in all applications has been dictated by their ability to complex proteins [4,6-9]. This allows them to form insoluble crosslinked structures which can inhibit functions of certain proteins. The phenolic nuclei make them excellent radical scavengers which lead to good antioxidant properties [10-15]. Their ability to tan leather is also based on their ability to complex with the collagen molecules on the surface of the hide [4,9].

The basic building blocks of tannin molecules are similar. However, differences are observed with regards to the number and location of hydroxyl groups along the oligomer chain and the degree of polymerisation. Although these differences exist, isolation of specific structures in tannin extracts has been a difficult task due to the inability of bulk techniques to isolate these differences. In order to understand the structure-property relationships of these molecules, specific oligomers have to be separated, isolated and identified such that various extracts may be compared. An understanding of these complex natural systems may well lead to further development in terms of their usage in various applications.

The aim of the research outlined in this thesis was to develop liquid chromatographic methods to analyse these heterogeneous oligomeric tannins extracted from wood. Since most studies have been performed on relatively low molar mass extracts, such as cacao bean extracts, this was used as the basis in order to develop the chromatographic method. By performing a comprehensive separation, molar masses may be assigned to specific oligomers and the presence of overlapping distributions will become clear.

2.2. Chemistry of tannins

Tannins are divided into two major groups, namely, condensed and hydrolysable tannins. Condensed tannins (also known as polyflavonoids or proanthocyanidins) are based on flavonoid units which undergo condensation and polymerisation reactions to form oligomers with varying degrees of polymerisation (DP). In nature these molecules are usually attached to their precursors, flavonoid analogs, carbohydrates as well as traces of amino and imino acids [9]. Hydrolysable tannins differ from condensed tannins; they are derivatives of gallic acid and are usually esterified to a carbohydrate core, mainly glucose [9,16-18]. However, these tannins often occur as complex mixtures of simple phenols (e.g. pyrogallol), gallic and digallic acids as well as esters of sugars and other structures (e.g. three dimensional networks) formed as a result of oxidative coupling and further esterification of the galloyl groups. The simplest of the hydrolysable tannins are gallotannins which are made up of polygalloyl esters of glucose such as pentgalloyl glucose (PGG) [4,18].

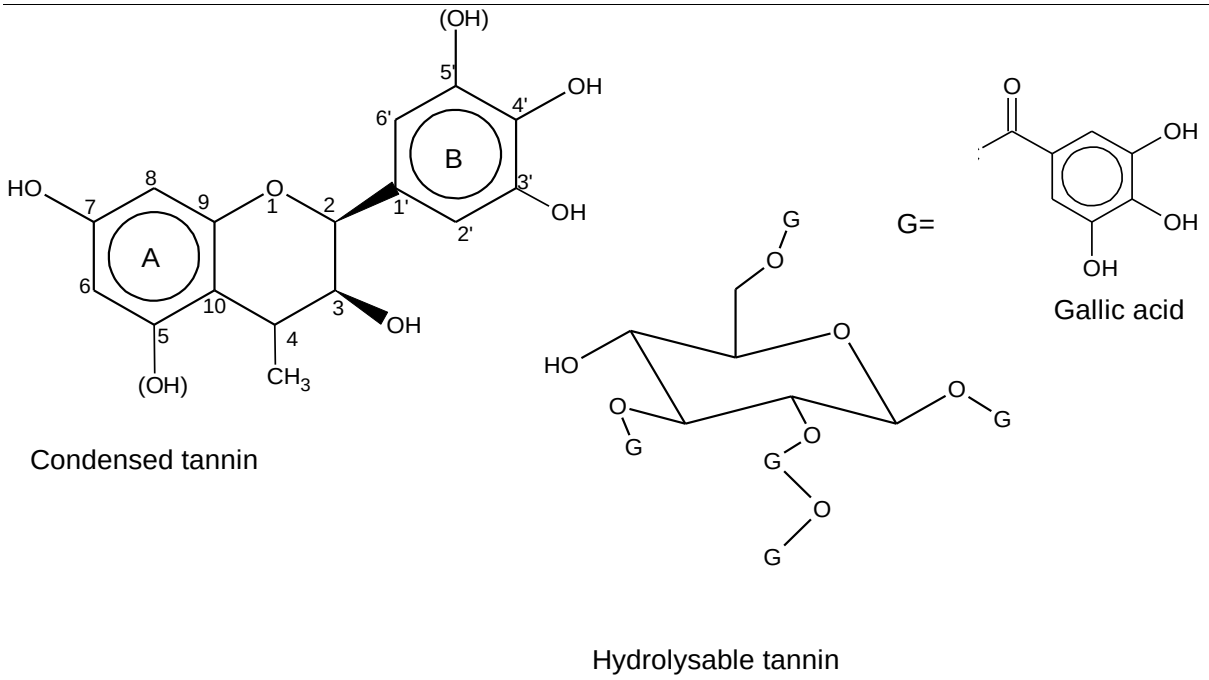


Figure 2.1: General structure of tannins

The use of tannin extracts is limited due to their availability and complex structures as well as the difficulty in producing high yield extracts. This being the case, they are still widely used in commercial applications such as leather tanning agents, flocculants, super-plasticisers, and for synthesis of wood adhesives [2-4]. In addition to these another use for tannins is in medical applications, they have been shown to have various health benefits such as radical scavenging properties which make them excellent antioxidants, antimicrobial effects, anti-tumour and antiviral abilities, anti-inflammatory abilities, treatment of heart and circulatory problems among others [5,6,19-24]. In addition, tannins have been discovered in many foods and beverages such as strawberries, apples, nuts, hops, apple juice, cocoa beans, wine and many others [6,23,25-31]. Tannin extracts have also been shown to vary in structure and degree of polymerisation depending on the part of the plant from which they have been extracted [5,26,32]. An example would be the condensed tannins extracted from grapes. The grape and grape seed extracts contain different types of condensed tannins [23,32,33]. The same is also applicable to extracts containing hydrolysable tannins. Extracts may be obtained from the galls and leaves of the sumac plant [7,18].

In all applications the function of tannins is dependent on the degree of polymerisation and the chemical structure [22,26,28,34]. Tannins are more widely used commercially as leather tanning agents than for the

synthesis of wood adhesives and medical applications combined [2]. Condensed tannins are preferred for synthesis of wood adhesives because they have more versatile structures which lead to superior reaction properties when compared to hydrolysable tannins which have reaction properties similar to those of simple phenols [2,9,16].

This work focuses mainly on the analysis of wood extracted condensed tannins obtained from different trees by various extraction methods. The analysis of hydrolysable tannins is also of interest although the study in this thesis was not as extensive. A brief overview indicating the differences in structure between the two groups will be outlined.

2.2.1. Hydrolysable tannins

Hydrolysable tannins are divided into two types, the first being those composed of mixtures of oligomeric simple phenols such as gallic and ellagic acid and the second consists of esters of sugars, mainly glucose with gallic and digallic acids [4,35,36]. The more complex structures of the latter may contain ellagic acid and are known as ellagitannins. Hydrolysable tannins are extracted from various sources, the most common being the bark of the chestnut (*Castanea sativa*) wood, tara (*Caesapina spinosa*) wood, oak (*Quercus infectoria*) galls, sumac (*Rhus Semialata*) galls, sumac leaves *(R. coriaria*, *R. typhina*) and chinese gall (from *Melaphis chinensis*). Galls are parasitic growths that form on the bark of some trees.

Gallotannins are the simplest of the hydrolysable tannins, they are formed by esterification of a sugar with gallic acid (3,4,5-trihydroxy benzoic acid), as previously mentioned, the most common sugar is glucose although other sugars such as hammamelose, shikimic acid, quinic acid and queratol have been detected in other plant species [18]. Gallic acid plays a significant role in the metabolism reactions and is formed via dehydrogenation of the 3-hydroshikimate [4,7], see **Figure 2.3**.

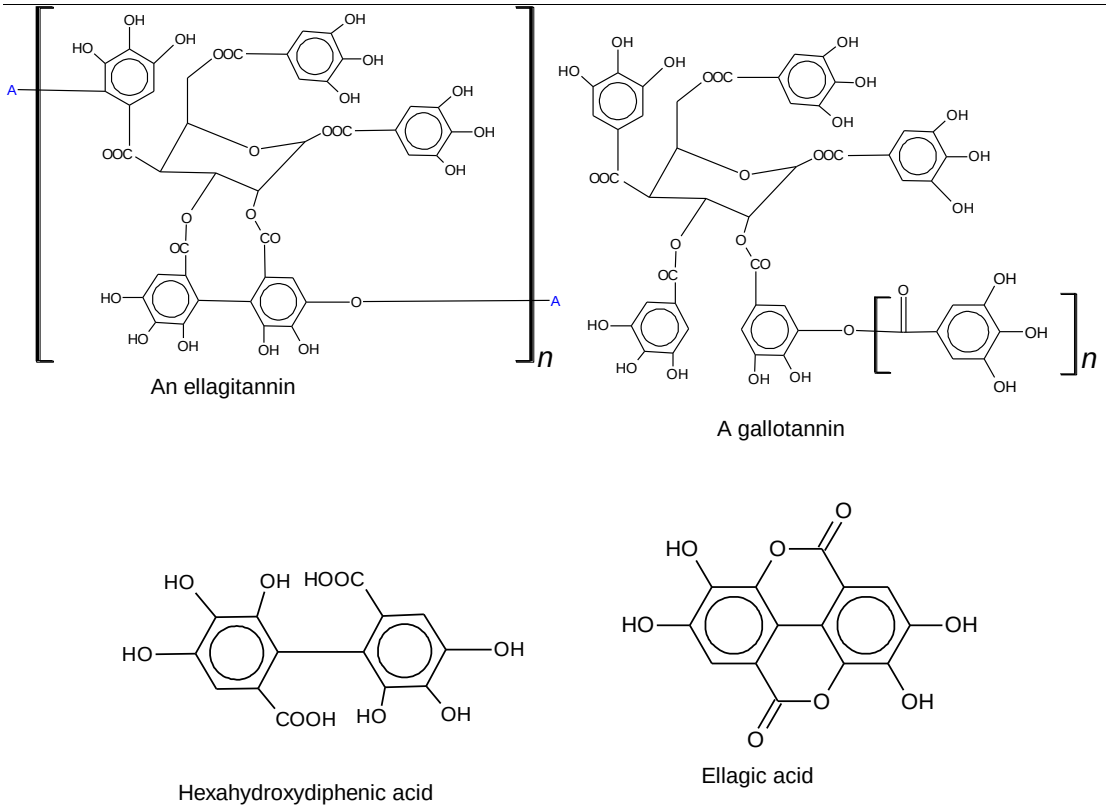


Figure 2.2: Structures of hydrolysable tannins and their precursor molecules.

The reaction of gallic acid with uridine diphosphate glucose (UDP-glucose) leads to the formation of β-glucogallin (β-1-O-galloyl-D-glucose), this precursor then undergoes several reactions which subsequently lead to the formation of pentagalloyl glucose (β-1,2,3,4,6-pentagalloyl-O-D-glucopyranose) or PGG; which is the primary reactant in the formation of many of the complex structures found in plant extracts [4,7,18].

The polygalloyl ester chains are formed either by meta- or para-depside bonds (as shown in **Figure 2.4**) via the phenolic hydroxyl groups and the oxidative coupling of these esters through formation of new C-C and C-O bonds yields hexahydroxydiphenol esters and their derivatives [4]. This oxidative coupling has been shown to be the reason for the structural variation in these molecules. The intermolecular C-O bond has been found in oligomeric structures. Sumac and oak galls were shown to contain simple gallotannins with galloyl esters that have up to 13 units and a core glucose [18,35]. The commercial tannic acid is merely a mixture of simple gallotannins for example sumac (*Rhus semialdata*) galls [18].

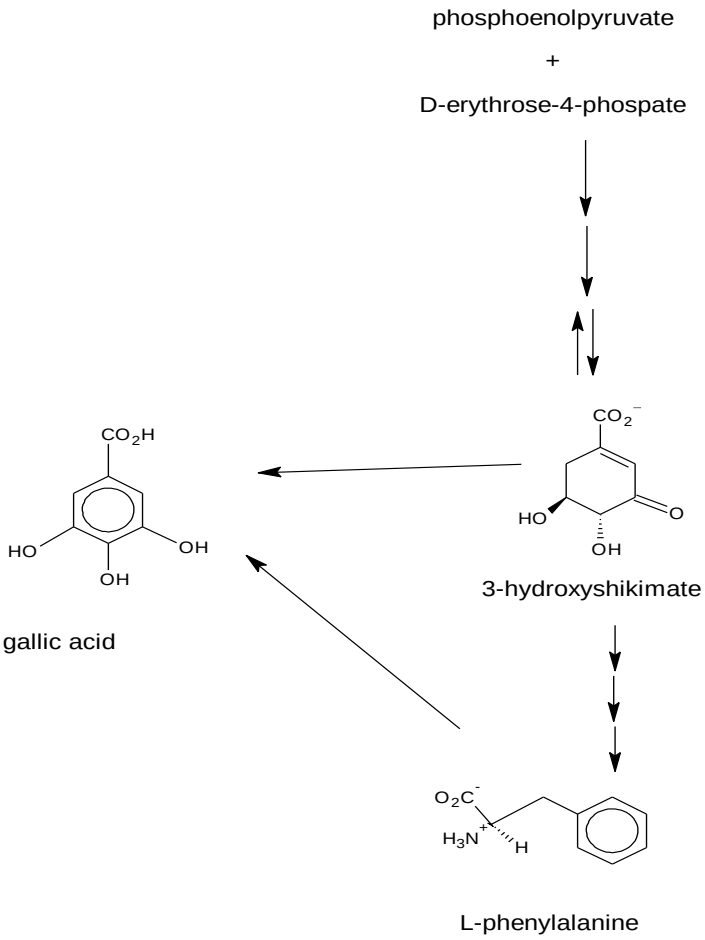


Figure 2.3: Proposed gallic acid biosynthesis via the dehydrogenation of the 3-hydroxyshikimate [4].

It is rare, however, in nature to find the simple gallotannins (**Figure 2.2**); instead natural extracts are mixtures of more complex structures such as ellagitannins. Ellagitannins are esters of hexahydroxydiphenic acid (HHDP) and are also formed by the oxidative coupling or dehydrogenation reactions of the galloyl groups of the gallic acid residues. HHDP will spontaneously lactonise in aqueous solutions in order to form ellagic acid (see **Figure 2.2**) [4,37]. Four chemical pathways are possible for the synthesis of the metabolites; the first pathway is the oxidative coupling as described above. The 'monomers' are formed by the C-C coupling whereas the oligomeric structures are formed by the intermolecular C-O coupling [4]. Depending on the form of the C-O bond, the structure that is formed will vary. Another type of reaction is whereby the gallic acid esters are formed with the ring-opening of the glucose pyranose ring. This reaction leads to formation of unique structures such as vescalin, vescalagin, castalin and castalagin, which have been detected in chestnut and oak tannins [4,18,38]. Castalagin and vescalagin are positional isomers, as are castalin and vescalin. Intramolecular coupling to form HHDP is common between C4/C6, C3/C6 and C1/C6 of PGG in its

$^1C_4$ conformation [4]. These different forms of hydrolysable tannins have been extracted and identified from the appropriate plant species.

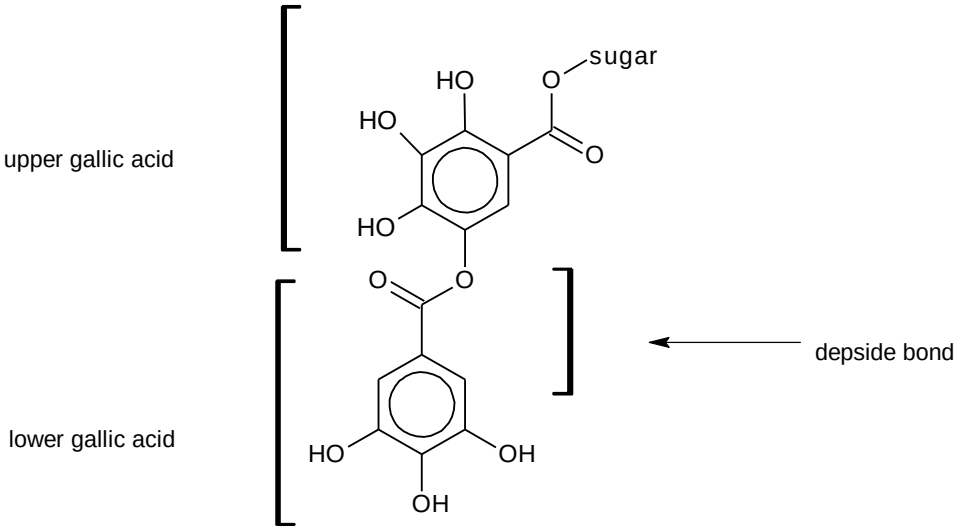


Figure 2.4: Depside bond as appears in some oligomeric hydrolysable tannins.

2.2.2. Condensed tannins

Tannin precursors are known as monoflavonoids which are simple phenolic structures namely, leucoanthocyanidins (flavan-3,4-diols), catechin (flavan-3-ols) and flavonols, flavones and coumaran-3-ones. Of these only flavan-3,4-diols and flavan-3-ols lead to tannin formation (**Figure 2.6**) [4,9,16,19,22]. Condensed tannins may also be referred to as proanthocyanidins and are extracted in large quantities commercially [2,17]. They are also found in a variety of food products such as apples, strawberries, other fruit and nuts as well as in cacao [4,26,32,39-43]. They have numerous health and nutritional benefits such as radical scavenging (enable them to act as antioxidants), protein binding, antimicrobial effects and anti-inflammatory properties [4,6,8,19]. The commercially extracted tannins are used in the manufacture of wood adhesives with varying gel times and different viscosities. These properties are influenced by the tannin structure and degree of polymerisation [3,9,11,16,44,45]. The structures of polyflavonoids differ in terms of the A-ring and B-ring structures, which may consist of different numbers of hydroxyl groups around the aromatic nuclei. Depending on this difference the structures are named differently: tannins composed of resorcinol A-rings and catechol B-rings are called fisetinidins and when composed mainly of resorcinol A-rings and pyrogallol B-rings they are known as robinetinidins. A phloroglucinol A–ring may combine with a

catechol or pyrogallol B-ring to form structures known as catechins and gallocatechins, respectively [45]. The above-mentioned precursors may combine in various ways to form more complex structures. These molecules tend to favour the linkage of either the C4-C8 or C4-C6 depending on the flavonoid structure (**Figure 2.6**) [3,9].

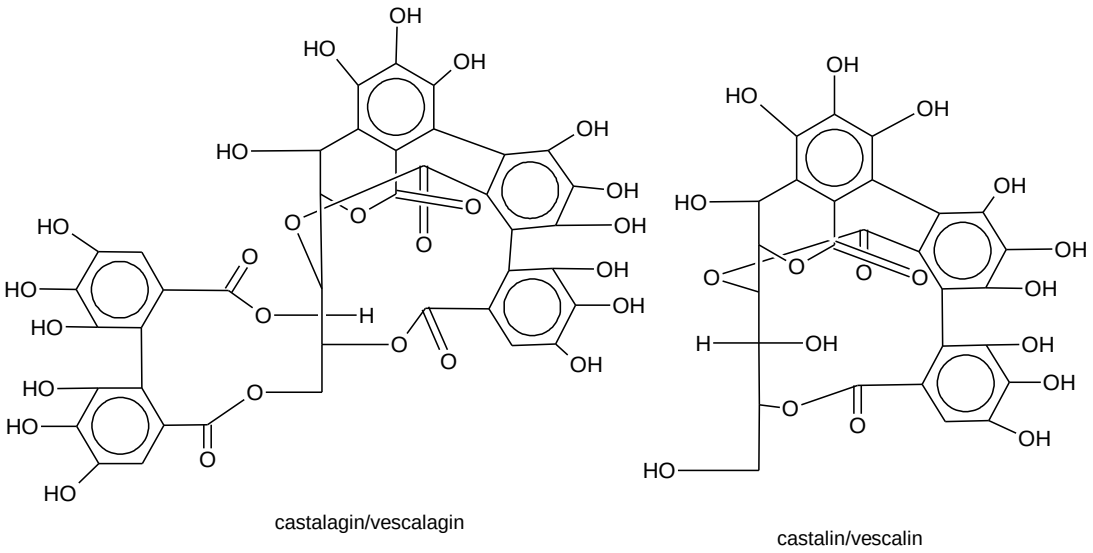
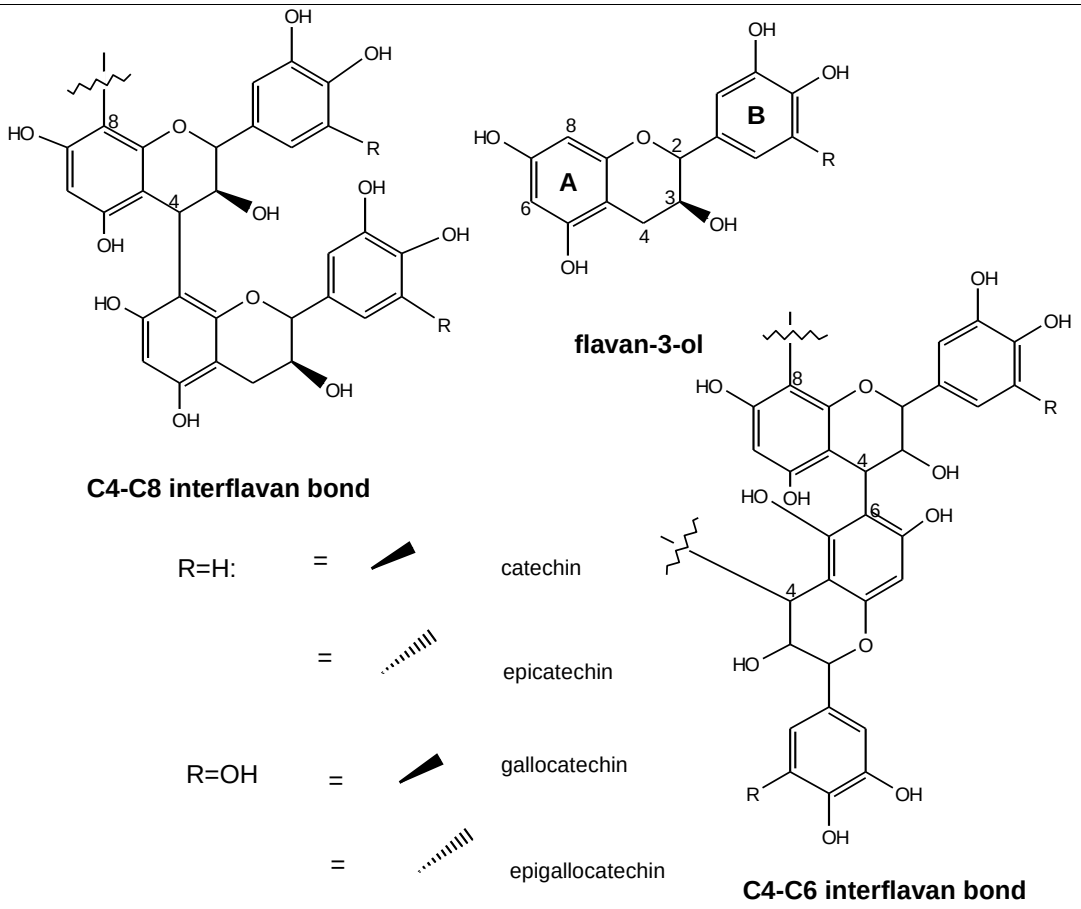


Figure 2.5: Structures of castalagin/vescalagin and castalin/vescalin: positional isomers formed from the ring-opening of glucose pyranose ring.

Although the exact nature of the reaction that leads to tannin formation is not well known, a well accepted explanation is that of autocondensation [9]. This reaction occurs due to the strong nucleophilic centres on the 6- and 8-positions of the A-ring, which are promoted by the metadisubstitution or -trisubstitution of the hydroxyl groups. The same type of stabilisation leads to formation of a benzyl carbonium at the C4 position, the positive charge is stabilised by the delocalisation on the vicinal aromatic ring. This same reaction occurs during curing of wood adhesives; under acidic or alkaline conditions the O1-C2 bond of the A-ring opens which leads to formation of a carbocation at the C2 position, this reacts with the free C6 or C8 of an adjacent unit on another chain and results in an increase in viscosity [11,46]. The strongest nucleophilic centres are found in catechin and gallocatechin which make them the most common structures found in tannin extracts.



R_1=OH, R_2=H: Procyanidin (PC)

R_{1+2}=OH: prodelphinidin (PD)

Figure 2.6: Structures of common condensed tannins and dimers showing the C4-C6 and C4-C8 interflavonoid bond.

The benefits of these compounds are many and have been extensively studied; and yet the structure-property relationships are still not well understood. HPLC has been the method of choice for the separation of proanthocyanidins for several years and yet it is still limited to the lower molecular mass compounds [34,47-50]. Wood extracted tannins have been shown by other techniques to contain polymeric proanthocyanidins [2,38,44,51-53].

The reaction between the condensed tannin and the curing agent is dependent on the structure of the tannin [46]. During a curing reaction the tannin molecules react with the curing agent to form a three dimensional network. Formaldehyde has been the curing agent of choice over the last decade, but this leads to problems of toxicity especially in indoor applications. As a result other less toxic curing agents have been proposed

such as hexamine and polyethyleneimine (PEI) [46,48,54,55]. The curing reaction is believed to occur through a process of condensation in alkaline or acidic conditions.

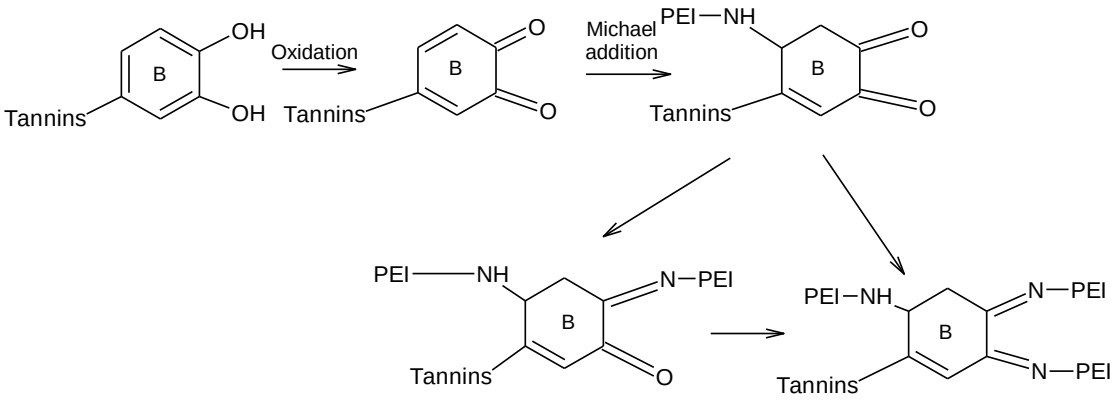


Figure 2.7: Scheme showing a possible reaction of tannins in the presence of polyethyleneimine (PEI) [46].

2.2.3. Extraction methods

Tannins are contained in various parts of the plant and depending on this the method of extraction may vary. The method of choice for the extraction of tannin is making use of an aqueous polar solvent. The most commonly used solvents for the extraction of the polyphenols from their various constituents are acetone and methanol-water mixtures (70:30 %v/v) [18,26,32,48,56-63]. Although this is the norm, other solvents and solvent combinations such as 60% methyl acetate or 70% aqueous ethanol may be used [31,33,64]. A low percentage of acid may be added in order to increase the stability against oxidation [32,60,63]. The drawback of these extraction methods is that they are only applicable to the lower oligomers and do not solubilise the polymeric fraction of the extract [42,65]. Therefore most of the analytical methods that have been developed have been based on these lower molecular weight fractions. In this thesis polymeric tannins are the ones with a DP higher than 14 units. It has been shown that some solvents have an affinity for certain structures and, therefore, will tend to exclude the structures that are not soluble and as a result the true form of the extract is not analysed but rather the soluble fraction [26,64,65]. In both proanthocyanidin and gallic/ellagic acid-containing plants the extraction by the aqueous solvent is often preceded by a

defatting step, whereby n-hexane or a similar non-polar solvent is used to remove lipids and chlorophyll (if present) [26,32,64,65].

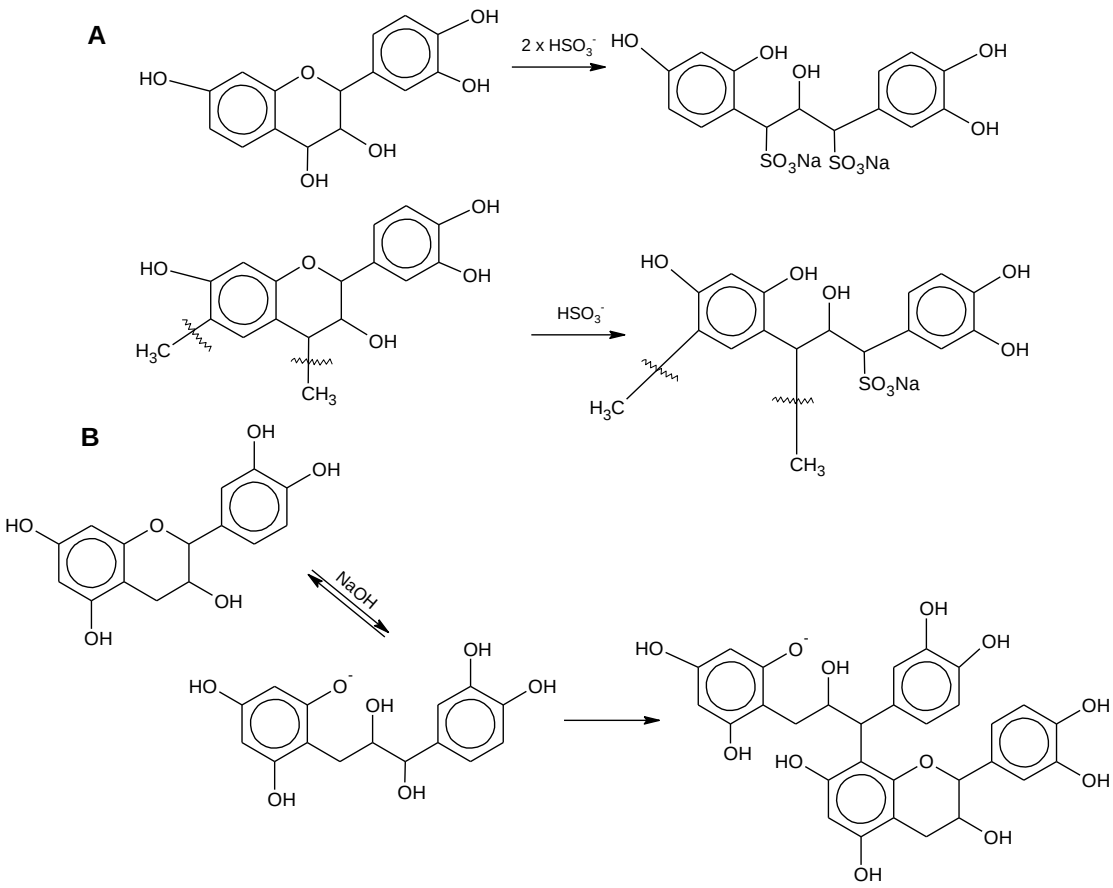


Figure 2.8: Reaction schemes showing the modification reactions of tannins. Scheme A shows the reaction of catechin with sodium metabisulphite ($NaHSO_3$), and Scheme B indicates the reaction of catechin with sodium hydroxide (NaOH)

Tannin extracts are often subjected to a further purification subsequent to extraction, Sephadex LH-20, Tosopearl HW-40 and SPE methods are often used for this purpose [4,31,66-69]. The former are usually used to separate the proanthocyanidins by degree of polymerisation, being able to separate monomeric, oligomeric and polymeric fractions of the plant extract [30,32,33,39,58,63,65]. In commercial extraction processes the tannins are extracted along with some precursors, sugars, amino and imino acids [9,16]. These extracts contain about 70% tannins [9]. This means that the lab extractions as described above do not represent the portion of tannins extracted industrially. In addition organic solvents are not permissible on an industrial scale due to waste disposal problems and pollution and, therefore, only water is used in the

extraction [3,9,47]. In tannin extraction for commercial use, the temperature may be elevated and the water medium may be sulphited or metabisulphited [3,9,16,47,70]. An important point to mention is that in industrial applications the pH is taken into consideration and the extraction medium may be acidified or alkalinated in order to improve reaction properties [3,71]. These industrial extraction methods often alter the structure of the tannin that is being extracted [36]. For example in the case where the extraction medium is sulphited, the hydroxyl groups are replaced by the sulphonic groups which vastly increase the solubility of the tannin (**Figure 2.8**) and may improve reactivity [9,16,70]. This is mainly the case in the production of wood adhesives, in which additives are added in order to improve solubility as well as reactivity of the tannin molecules. In some instances NaOH is added, this increases the pH and thus promotes the reaction of the tannins with certain additives like formaldehyde or hexamine which are added in the production of wood adhesives [2,3,9,55,71].

For applications in wood adhesives, quebracho, mimosa and pine extracts are the ones that are used to a greater extent. In leather tanning which accounts for the higher percentage in the use of tannins, hydrolysable tannins such as sumac, tara, turkey gall, and chestnut are mainly used, however, for tougher leather condensed tannins used for wood-adhesive applications are also employed [2]. One of the most well studied aspects of the tannins is their health benefits as mentioned earlier, thus, it would be preferred if the extraction would result in extracts that contain only the tannin molecules.

2.3. High Performance Liquid Chromatography (HPLC): Principles of separation

This work focuses on the characterisation of oligomeric tannin structures. Tannin extracts contain various constituents which make it difficult to fully characterise the materials. In addition these are natural extracts and their bio-synthetic method is still not well understood which adds to the complexity of the problem. Tannins will, therefore, have distributions in chemical composition and molar mass, and the number of hydroxyl groups per repeat unit varies. Due to the complexity of the samples, one mode of separation will not provide sufficient information, therefore, to fully understand the system a combination of several techniques is required. Tannins used in industrial applications have been analysed extensively by bulk techniques such as MALDI-TOF mass spectrometry and NMR [2,35,36,38,44,45,72,73]. However, limited information is available with regards to the liquid chromatographic separations carried out [52]. This could be due to their

higher molar mass and the added complexity introduced by the extraction methods. Extensive reviews on the applications and limitations of different HPLC techniques for the analysis of synthetic polymers have been covered elsewhere [74,75]. In these reviews the focus is mainly on the analysis of synthetic polymers and whereby the important aspects of chromatography are covered.

An HPLC separation relies on the strength of interactions of the analyte with the stationary and/or mobile phase. Depending on the strength of these interactions a separation by size or chemical composition may be obtained. The distribution of the analyte between the stationary and the mobile phases is governed by thermodynamic effects as well as the chromatographic distribution coefficient (K_D). The distribution coefficient will, therefore, determine the order of elution of analytes. The distribution coefficient can be described as the ratio of the concentrations of the analyte in the stationary phase and the mobile phase (**Equation 1**).

$$K_D = \frac{[Analyte]_s}{[Analyte]_m} \quad \textbf{(1)}$$

In liquid chromatography porous particles are often used as the stationary phase. The pores have limited dimensions and depending on its molecular mass the analyte molecule can either fully or partially penetrate the pores of the stationary phase where the active sites are normally located. Two modes of chromatography are possible here, steric exclusion and adsorption/partition. The retention volume (V_r) for a liquid chromatographic separation of analyte molecules can be described as:

$$V_r = V_i + V_{stat} K_D \quad \textbf{(2)}$$

V_i is the interstitial volume, V_{stat} is the total 'stationary' volume in the column and it is composed of the pore volume (V_p) and the volume of the stationary phase (V_s) that is active in adsorption processes. Therefore the equation may be expanded to:

$$V_r = V_i + K_{\text{sec}} V_p + K_{ads} V_s \quad \textbf{(3)}$$

K_{sec} represents the distribution coefficient of steric exclusion and K_{ads} the coefficient for adsorption. K_D is an equilibrium constant related to the change in Gibbs free energy (ΔG) which is related to the partitioning of the analyte molecule between the stationary and mobile phases. Therefore, the change in Gibbs free energy is also affected by the steric and adsorption effects.

$$\Delta G = \Delta H - T\Delta S = -RT \ln K_D \qquad \textbf{(4)}$$

The active sites of the stationary phase for the chromatographic separation of analyte molecules may be located on the surface of the stationary phase or inside the pores [74,76]. A thin adsorption layer with width, *a*, is formed on the surface of the solid stationary phase and for low molar mass compounds chromatographic interactions are centred here [74]. This volume (V_a) consists of a small fraction of the pore volume (V_p) such that $V_a << V_p$. Thus, **Equation 2** still holds for small molecules. It is possible to suppress enthalpic interactions (ΔH = 0) with the stationary phase by using a thermodynamically strong solvent, so that K_{ads} = 0 and the separation is governed by entropic exclusion effects [77,78]. This chromatographic situation describes the *ideal* size exclusion (SEC) mode of separation and K_{sec} is between 0 and 1; when it is zero the analyte molecules are too large to penetrate the pores and are said to have reached the exclusion limit of the column. If however K_{sec} is 1, this means that all the analyte molecules are able to penetrate all pores of the stationary phase and leads to total permeation and thus no separation. The separation in SEC mode takes place when $0<K_{SEC}<1$, therefore the large molecules will elute first and the small molecules will elute last i.e. separation is by molar mass. If a poor solvent is used or the temperature is altered then the enthalpic interactions are no longer suppressed and adsorptive interactions of the analyte and the stationary phase occurs. This mode is known as the liquid adsorption chromatography (LAC) mode. In the ideal case it means there will be no steric interactions and therefore ΔS = 0. The *ideal* LAC mode can only occur if the pores are large enough for all the molecules to penetrate or small enough that all the molecules are excluded. This is hardly ever the case in real systems since the analyte molecules have various sizes and, therefore, will interact differently with the single size pores of the stationary phase. In addition the interaction of the analyte molecules from the mobile phase involves conformational changes in order to access the functional groups of the stationary phase [76]. Some molecules will enter the pores and some will be excluded, therefore, both enthalpic and entropic effects are occurring at the same time. In the LAC mode

both the entropy and enthalpy term affect the separation and thus the elution volume can be described such that the smaller molecules are eluted first followed by the larger ones.

In SEC mode usually isocratic and isothermal conditions are used to achieve separation. In LAC however, isocratic conditions are insufficient since both enthalpic and entropic interactions are occurring simultaneously and, therefore, a specific stationary phase-solvent combination has to be used in order to facilitate the required type of separation [76]. Frequently gradient elution is applied, and in the case of natural and synthetic polymers a binary solvent system is often used. Using solvents with varying strengths, the separation can be improved. A weak solvent is often combined with a stronger solvent to influence the interaction of the analyte molecules with the stationary phase and thus their elution from the column [79]. In the case of binary solvents which contain a certain amount of the stronger solvent 'B'; the eluent composition on the surface of the solid (Φ_{surf}) may be different to the total eluent composition Φ. This is due to the fact that solvent 'B' molecules preferably adsorb to the surface of the stationary phase. In addition when analyte molecules or monomer molecules from a chain adhere to the surface of the stationary phase, they replace the solvent molecules near the surface, and thus adsorb to the surface [74,78]. Usually $\Phi_{surf} > \Phi$ because the solvent 'B' will more strongly interact with the functional groups on the surface of the solid pore surface and this is dependent on the interaction energies of the components of the eluent and the stationary phase [78,80]. LAC is governed mainly by interaction and thus $\Delta G < 0$, and thus solvent 'B' will preferentially adsorb to the stationary phase.

In LAC there are different modes that can be used; these modes describe the various combinations of mobile and stationary phases.

2.3.1. Normal Phase Liquid Chromatography (NP-LC)

To achieve a normal phase separation a polar stationary phase is used in combination with a non-polar mobile phase under isocratic conditions. As mentioned earlier, isocratic elution is often insufficient to achieve separation. In gradient elution the separation is started with a non-polar solvent and gradually going to the polar one. Silica based stationary phases can be modified to adjust the polarity, for normal phase OH<NO_2<NH_2<Diol<CN, with silica being the most polar. Separation in this case is by increasing polarity.

The most commonly used polar stationary phase in the analysis of natural extracts is silica [7,26,32,39,42,58].

2.3.2. Reversed Phase Liquid Chromatography (RP-LC)

The elution order of reversed phase chromatography (RP) is reversed when compared to normal phase. In this case the gradient is begun with a polar solvent and going to a less polar one. The polarity of the RP stationary phase may be adjusted as well. Silica grafted with large non-polar aliphatic groups is used as stationary phase; the length of the aliphatic groups determines the polarity. C18 is the least polar and the most commonly used, C8>C5>C4>C3>C2>C1 with decreasing lipophilicity [77,81].

2.3.3. Hydrophilic Interaction Liquid Chromatography (HILIC)

Hydrophilic Interaction Chromatography (HILIC) is another mode of NP-LC which makes use of hydrophilic neutral stationary phases in combination with aqueous-organic mobile phases [6,82-87]. The stationary phases such as silica, silica gels modified with many polar functionalities such as amide, derivatives of poly(succinimide), sulfoalkylbetaineamino, cyano, diol, and polar modifications of silica with other groups such as carbamoyl groups are used for HILIC and diol-containing polymers can also be used [85,88,89]. The elution order is opposite to that obtained with RP-LC, in that retention increases with an increase in water content [6,42,84]. A water layer is formed on the surface of the stationary phase because of the hydrophilic groups versus the hydrophobic mobile phase and partitioning of the analyte molecules occurs between the mobile phase and the water-layer [82,84,85,90].

The separation in HILIC combines different interactions including electrostatic mechanisms. Hydrogen donor interactions and ion-exchange may occur with some analytes [90,91]. HILIC has been successfully applied to analyse peptides, carbohydrates, pharmaceutical drugs, plant extracts and oligomeric proanthocyanidins [42,83,84,87,91,92]. The mechanism of HILIC is described in detail in a review by Hemstrom and Irgum [82] and the stationary phases were covered at length by Ikegami [82,89]. HILIC was also shown to be superior to silica based separations in the analysis of oligomeric proanthocyanidins [42,83].

2.3.4. Detection

2.3.4.1. Selective detectors

Selective detectors are spectroscopic detectors that monitor a single functional group in the molecule. There are a number of selective detectors that are available, namely, infrared (FTIR), electrochemical, ultraviolet, fluorescence and ^{1}H-NMR. These detectors are able to provide valuable specific information about the structure of the analyte but they have the drawback that the analyte molecules must consist of a specific group(s) and thus limit their use [93]. The most commonly used detector in the analysis of plant extracts is UV detection and for polyphenols the second major one is fluorescence detection [7,34,39,42,49,50,94]. In the analysis of polyphenols this drawback is not important since all the molecules will be analysed. However, in the case whereby other structures may be present, they will not be detected if they are not UV absorbing. An advantage of using UV as a detector is that it is easily employed for quantification. The combination of UV and other detectors may provide valuable information about the structure and chemical composition of the analyte molecules. ^{1}H NMR can now be directly coupled to an LC separation. The application of LC-NMR is very limited especially for the analysis of polyphenols [95]. The most important information in structural elucidation of natural products is obtained in the ^{13}C NMR spectrum. In order to obtain this spectrum higher sample concentrations are required. In order to achieve an optimum LC-NMR coupling high sample loading is required and thus this leads to the problem of solubility in the case of polyphenols. In addition, the higher molar masses of these extracts make it difficult to achieve good detection, due to low sensitivity inherent in the technique [95]. Structural assignments of the ^{1}H signals become even more complicated with these types of structures. However, this combination has been shown to provide valuable information with regards to identifying the components of plant extracts and identifying new compounds [7,93,95,96]. Most of the applications have been focused on low molar mass compounds but some attempts were made to analyse relatively higher molar mass compounds [93,96,97]. LC-NMR in combination with other techniques such as HPLC-UV-MS may give valuable information towards structure elucidation [93]. The merits and drawbacks of LC-NMR in application to polyphenols have discussed in detail elsewhere [97]. Off-line NMR has been employed and provides important information with regards to structure and is necessary in order to obtain a

full structural assignment [7,98]. ^{13}C NMR is a well established technique for the analysis of crude tannin extracts and if these extracts can be separated by LC in order to get more homogenous fractions a more detailed look into structure may provide important information on the isomeric structures [98]. A factor that limits the use of ^{1}H NMR for the analysis of high molar mass proanthocyanidin extracts is the complexity of the resulting spectra but if homogenous fractions can be obtained it will simplify the problem and thus more information may obtained with regards to isomeric structures.

2.3.4.2. Universal detectors

The two most common universal detectors used in liquid chromatography are refractive index (RI) and evaporative light scattering (ELSD) detectors. In the case of the RI detector the change in the refractive index of the mobile phase caused by the dissolved molecules is monitored [81,99]. The use of this detector is however limited by the fact that it may not be used with gradients and this is the most common form of analysis used for the separation of polyphenols, but it can be used in SEC analyses. As an alternative the ELSD is used since it measures any non-volatile components in the mobile phase; as the solvent leaves the column it is nebulised and then the solvent is evaporated which leaves solid particles. The particles are allowed to pass through a light beam being carried by a gas and the scattered light is measured and gives the signal. The ELSD response is non-linear and it is more complicated in relation to concentration and thus this may be one of the reasons for its limited use in the analysis of polyphenols [99]. Calibration of this detector may be carried out if the proper relationships are used; it may give concentration information [99]. For low UV absorbing compounds this is a very good alternative. In the case of natural extracts this detector has only been used to analyse low molar mass compounds such as flavonoids, terpene lactones and low molar mass polyphenols [14,68,100-102]. Lokvam and Kursar combined ELSD detection with electrospray ionisation mass spectrometry (ESI-MS) in order to analyse derivatised tannin oligomers from *Inga umbellifera* (tropical forest tree) leaves, this was used to determine the DP of the oligomers as observed in the ELSD signals [102].

2.3.4.3. Molar mass sensitive detectors

Since the tannin extracts are quite low in molar mass as compared to real polymers, typical molar mass sensitive detectors such as the viscometer detector or the light scattering detector cannot be used.

The MALDI-TOF technique was developed by Tanaka (1988) and was developed by Karas and Hillenkamp for biopolymers, MALDI-TOF is being widely applied in different fields in order to obtain molar mass and structure information [103-105]. This technique is able to analyse polymer molecules intact up to high molar masses. A dilute solution of the analyte is mixed with a light absorbing matrix. Molecules used as MALDI matrices are chromophores which are usually aromatic organic acids. In application to tannin extracts the most commonly used matrices are 2,5-dihydroxy benzoic acid (DHB) and α-cyano-4-hydroxycinnamic acid (HCCA) since they produce the best spectra [73,106]. A small amount of the matrix/analyte solution is placed on a sample target, then allowed to dry and crystallise. The target is then introduced to the ion source of the instrument and then a laser irradiates the target. The laser provides energy to the matrix and analyte molecules causing them to be transferred to the gas phase. Various mechanisms have been suggested for ion formation following the laser irradiation. The most widely accepted mechanism for primary ion formation is whereby two or more matrix ions are excited and then combine their energy to form a high energy matrix molecule or matrix radical ion and then this charge is transferred to the analyte molecules [107]. A salt, commonly alkali halogenides are often added in order to assist the ionisation process. The ions formed during irradiation are accelerated by an electric potential to a fixed kinetic energy and directed into a field-free flight tube (drift tube) [103,105,107]. The flight tube is where the accelerated ions are separated according to their mass-to-charge ratio (m/z) and subsequently reach the detector [107-111].

In the analysis of both condensed and hydrolysable tannins MALDI-TOF was used very successfully to determine the chemical composition of complex plant extracts [2,6,35,36,38,44,65]. The analysis of tannins using MALDI-TOF is a relatively new technique and provides additional information on the structure and molar mass distributions in a single experiment. Pasch et al. applied MALDI-TOF to analyse polymeric tannins, the analysis of quebracho and mimosa extracts showed the difference in composition of these two similar wood extracts [38]. Quebracho was shown to contain mainly profisetinidins which lead to the formation of linear structures whereas mimosa is predominantly composed of prorobinetinidins [38]. In the

analysis of the chestnut extract, a hydrolysable tannin, the presence of vescalin/castalin and vescalagin/castalagin was conclusively determined [2,36]. MALDI-TOF has been successfully applied offline to analyse fractions obtained from a chromatographic separation, information obtained from this technique is very valuable since it can reach very high molar masses [65,73,81,107]. ESI-MS tends to form multiply charged molecules, MALDI-TOF spectra provide the advantage that only singly charged molecules are formed and thus making the spectra easier to interpret.

A new technique for the analysis of tannin monomer sequences is post source decay (PSD) fragmentation whereby a specific ion is selected from a MALDI-TOF spectrum and is subjected to higher laser intensities [73,110]. The high laser intensity results in fragmentation of the 'mother' ion and these fragments are detected. Ion dissociation is induced by the excess of internal energy that ions gain from the laser during the ionisation step. When the ions fragment in the field free region the process is termed PSD [73,112]. Behrens et al. was the first to show the applicability of this method to tannins and condensed tannins from lime and spruce [73]. Another method that performs a similar type of analysis is collision induced dissociation (CID) whereby the precursor ion is selected from the first TOF analyser and introduced into a collision cell whereby it collides with inert gas molecules. The fragments formed are then reflected and analysed in the second TOF analyser [112-114]. Although this technique has not been used for the analysis of tannins it has been shown to be applicable to biopolymers as well as synthetic polymers and thus shows great promise for analysis of tannins [112,113].

Unlike MALDI-TOF and FAB, ESI-MS does not make use of a matrix, although it is also a soft ionisation technique [110]. In ESI-MS a dilute analyte solution is introduced at a constant flow rate to a smaller capillary or needle kept at high voltage (0.5-5kV) [110]. Highly charged molecules are formed upon exit from the needle (Taylor cone) due to the high potential present [115]. The solution is sprayed and the solvent evaporates which leads to the formation of a charge plume. The highly charged droplets combine and are detected [110,115]. Due to the ability of ESI-MS to form multiply charged ions high molecular mass molecules can be analysed if they have a *m/z* in the range of 500 to 2500 Da [115]. ESI-MS has been widely applied in structure elucidation of oligomeric tannins [5]. On-line LC-ESI-MS is the popular choice for the analysis of tannins since it is able to analyse intact molecules [32,39,42,116]. Guyot et al. and other workers have successfully applied LC-ESI-MS to analyse condensed tannin structures [26,32,39,42]. Nunez et al.

went as far as being able to distinguish between galloyted and non-galloyted condensed tannins extracted from grape seeds and Hammerstone *et al.* showed the wide applicability of the LC-ESI-MS method to various tannin extracts [32,39,63]. Zywicki et al. applied this method to analyse wattle, a condensed tannin extract and chestnut, a hydrolysable tannin extract and was able to conclusively determine the structures of these tannins [52].

2.4. Analysis of tannin chemical structure

2.4.1. Bulk techniques for analysis of oligomeric tannins

Many attempts have been made to find good analytical methods for tannins. Due to the complexity of these molecules this has proven to be a challenging task. Bulk analytical methods applied to tannin extracts greatly depend on the type of information that is required. The most widely used methods are calorimetric assays which can be used either to distinguish between hydrolysable and condensed tannins or to obtain specific information about structure. The acid-butanol assay may be used to determine the amount condensed tannins; this assay will only react with the condensed tannins. The reaction is shown below (**Figure 2.9**), it occurs by cleavage of the interflavonoid bond and produces anthocyanidins which are red coloured compounds [4,8,37,117]. The vanillin assay is similar to the butanol assay, the vanillin reacts with the polyflavonoid molecule in the presence of an acid such as sulphuric acid to form a red coloured compound. Unlike the butanol assay the vanillin assay is not specific to condensed tannins; it will react with any phenolic material present in the solution [1,8].

Other common methods for the analysis of condensed tannins are acid-catalysed degradation reactions whereby the interflavan bond of the condensed tannins is broken to form chemical derivatives of the tannin molecules. The most common of these reactions is thiolysis which occurs when condensed tannins are heated in the presence of acid and benzyl mercaptan [8,19,22,117,118]. On completion of the reaction the terminal unit appears as a flavonoid whereas the internal units on the chain form benzyl thioethers. A similar type of reaction makes use of phloroglucinol which does not form derivatives. However, the thiolysis reaction

is favoured because it produces higher yields and has been sufficiently developed and thus reliable. The acid-catalysed reactions are shown in **Figure 2.10.**

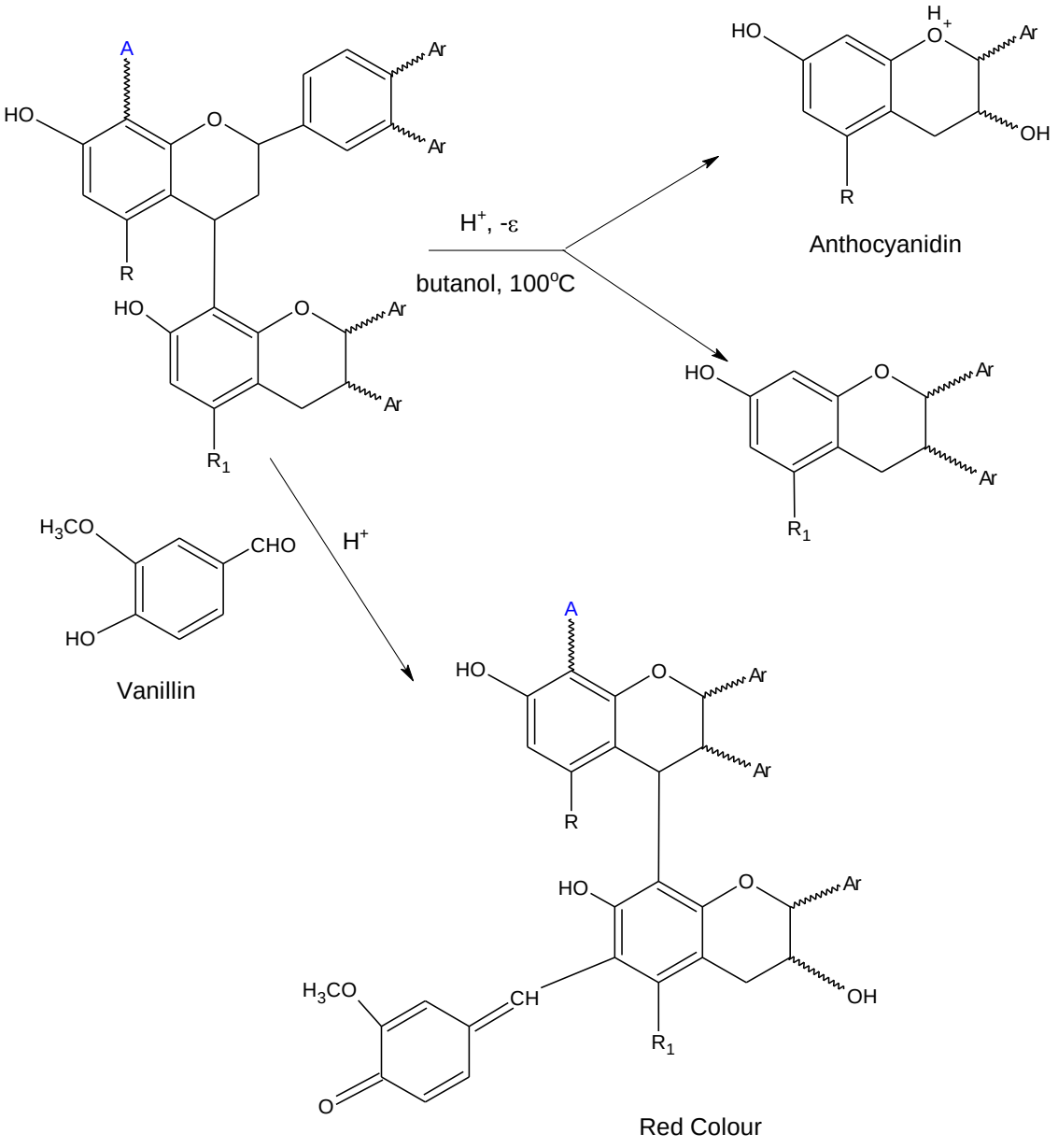


Figure 2.9: Chemistry of calorimetric assays used for condensed tannins. A-butanol assay and B-vanillin assay.

The Bate-Smith assay is specific for gallo- and ellagitannins; this reaction occurs with a KIO_3 reagent and yields a pink product but is not suitable for the analysis of complex mixtures. The rhodamine assay is very specific for gallotannins; the rhodamine reagent will only react with gallic acid. The assay is applied by analysing the amount of gallic acid before and after hydrolysis of gallo- and ellagitannins [7]. The $NaNO_2$

reagent is specific for ellagic acid and thus this means that the assays developed for hydrolysable tannins may underestimate the actual amount of tannins present (**Figure 2.11**) [7,119].

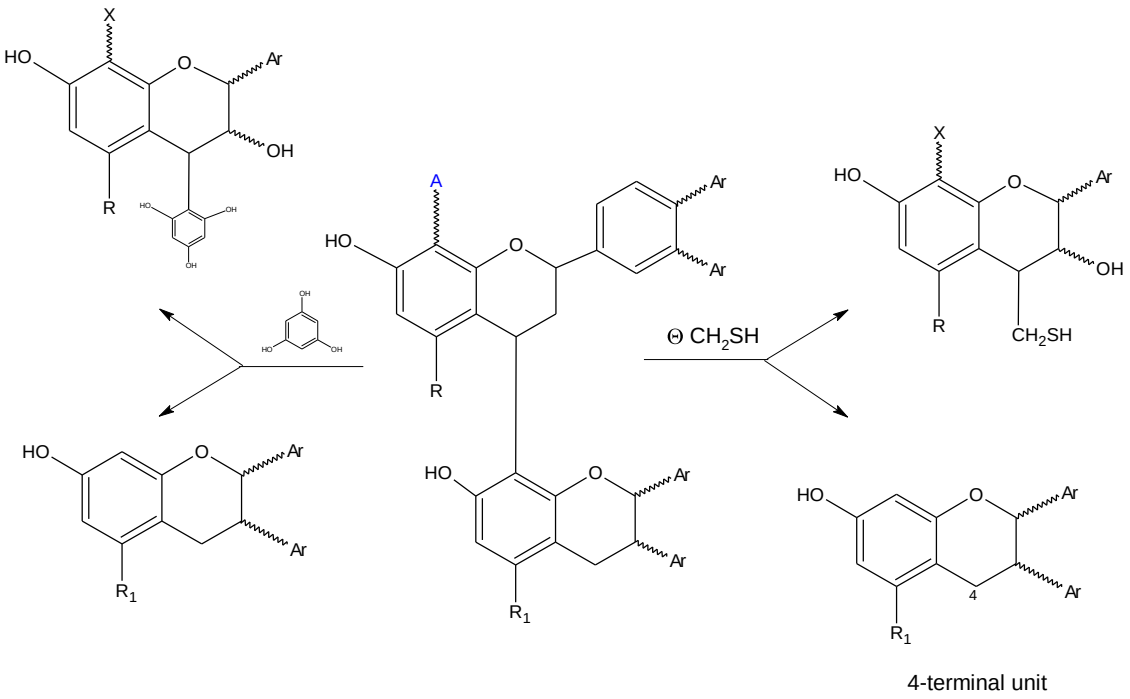


Figure 2.10: Chemistry of acid-catalysed reactions for thiolysis and phloroglucinol reaction.

Spectrometric methods have also been used successfully to identify and quantify the amount of tannins present in an extract. Both ^{1}H and ^{13}C nuclear magnetic resonance (NMR) spectra have been widely used to determine the chemical structure and average degree of polymerisation (DP) of extracts [45,120]. Soft ionisation mass spectrometric methods such as fast atom bombardment (FAB), ESI and MALDI-TOF mass spectrometry have been widely applied to elucidate the structure of complex mixtures of tannins. For analysis of tannins, ESI and MALDI-TOF are more widely used [110].

^{13}C NMR has been used extensively in the study of tannin structures. The following information on the structure of the proanthocyanidins can be acquired; (i) the ratio of procyanidin (PC) to prodelphinidin (PD) units, (ii) the stereochemistry of the heterocyclic ring of the monomer units, (iii) the number average molar mass (M_n), (iv) the extent of the heterocycle opening. All this information can be obtained from a single experiment. ^{13}C NMR is one of the most rapid methods for determining the required tannin characteristics. Czochanska *et al*. developed such a method using d_6-acetone-D_2O solutions of the tannin extracts [121].

Pizzi *et al*. developed a similar method using D_2O as the solvent with the extracts diluted 40% mass/mass with D_2O in 1:4 mass proportions [45,121]. The various carbons give distinct signals with minimal overlap and thus various extracts from different plants or chemically modified samples can be compared. The peaks are assigned according to the assignments proposed by Czochanska *et al* [121]. ^{13}C NMR can also be applied to determine the presence of various tannin ring structures but fails however to isolate specific structures and to indicate the relative amounts in a particular extract [45]. A relative method may be used to determine the mentioned proportions making use of the C5 and C7 signals on the tannin ring structure (shown in **Figure 2.12**) that appear at 156-157 ppm in the spectrum. The degree of branching can also be indicated by the proportion of the C4-C8 linkage. In the extracts where there are some carbohydrates present, peaks appear at 60-90 ppm and these overlap with the C2 and C3 signals that appear in the monomeric/pure catechin. ^{1}H NMR is also used to a large extent to determine the structures of low molar mass extracts of tannins [45].

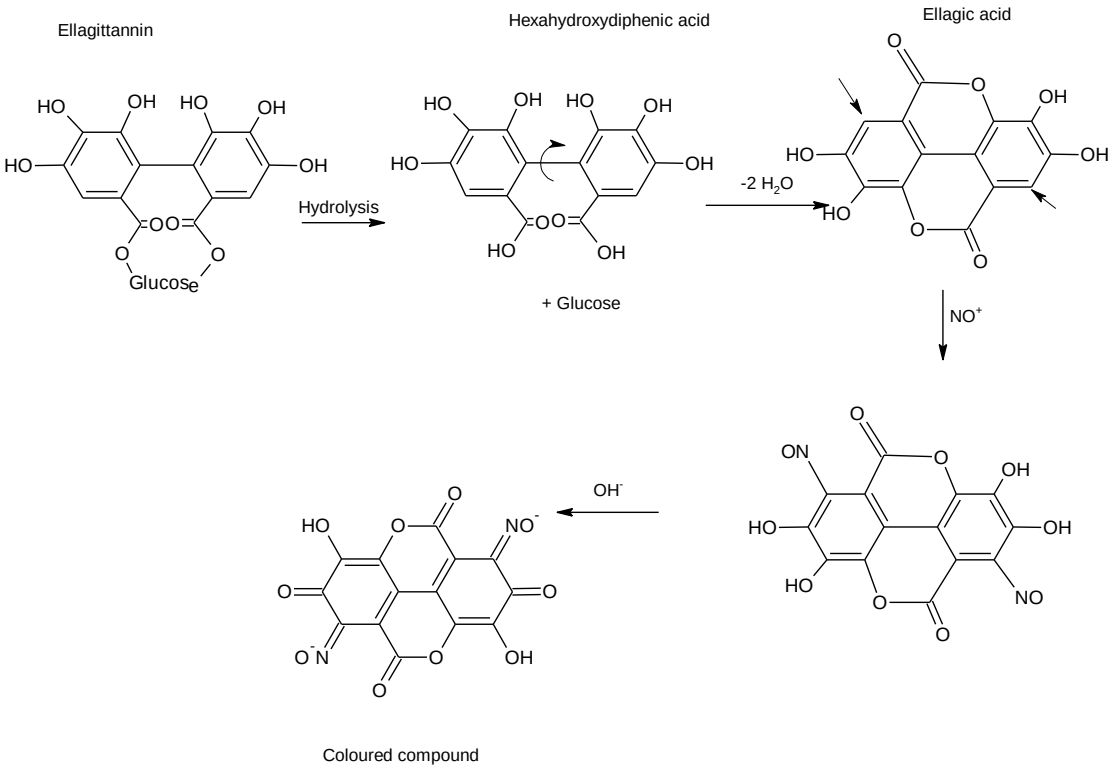


Figure 2.11: Chemistry of $NaNO_2$ assay for the analysis of ellagitannins.

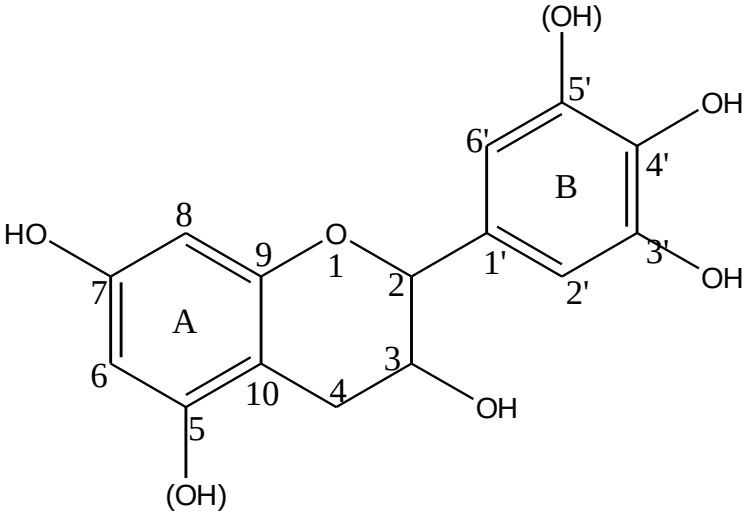


Figure 2.12: Generalised condensed tannin structure with labelling used to identify peaks in NMR.

2.4.2. High Performance Liquid Chromatography (HPLC): Separation of tannins

Tannins have wide chemical composition and molar mass distributions. Thus far though there is no chromatographic method that is robust and selective enough to separate polymeric tannins. This is due to the fact that both condensed and hydrolysable tannins have complex structures that are further complicated when the molar mass is increased. In condensed tannins it has been observed that an increase in molar mass causes an increase in complexity because the possibility of isomer formation increases [122]. The use of one chromatographic method has proven insufficient to completely separate these molecules. In hydrolysable tannins other types of structures such vescalin and vescalagin are found in other plant extracts and may co-elute with the other oligomeric gallic- and ellagitannins.

Normal phase columns with silica as stationary phase are used to separate tannins by the degree of polymerisation (molar mass) and RP-C18-based stationary phases are used to separate molecules by chemical composition. Hagerman et al. used a method adapted from Wilson to separate galloyl esters from fire weed flowers by their molecular mass on a silica column with hexane and a combination of tetrahydrofuran and methanol [1].

Reversed phase separation of hydrolysable tannins was performed by Tanaka (1984) and showed that the resulting chromatograms were very complex. Separation was achieved by chemical composition superimposed on a size separation [94]. Therefore, it can be surmised that normal phase chromatography achieves a size separation whereas reverse phase separates by chemical composition [19]. Although this is the case, complete separation is not achieved. At higher elution volumes a 'hump' appears in the chromatogram which represents the polymeric fraction of the extract. Attempts have not been made to resolve this hump in the chromatogram.

In the case of condensed tannins a great amount of work has been done to achieve good separations using liquid chromatography. The separation of cacao bean extract by normal phase LC was achieved by Rigaud et al. who showed that separation is by degree of polymerisation for the catechin/epicatechin series of oligomers [57]. Although the separation was good, at higher elution volumes the oligomers appeared as a large hump. Hammerstone et al. utilised a silica stationary phase in combination with dichloromethane-methanol-acetic acid and water; the application of a gradient with this ternary mobile phase enabled separation of the cacao bean extract up to tetradecamers [39]. The same method may be used to separate condensed tannins contained in other extracts [32].

Normal phase separation is governed by adsorption and, therefore, this leads the separation of the polymeric fraction of the extract to be incomplete. In addition it has been noticed that as the degree of polymerisation increases so does the number of isomers and the amount of substitution. This leads to co-elution even in normal phase LC which predominantly separates according to DP [26,31,57]. Catechin and epicatechin molecules in some extracts may be esterified with gallic acid, these galloyl groups affect the separation of the molecules since they may co-elute with oligomers that have the same molar mass but different chemical compositions [26]. Nunez et al. investigated this phenomenon in grape and grape seed extracts and was able to identify the galloyted molecules that were co-eluting with other procyanidin oligomers [63]. As far as is known there has been no direct separation of higher molar mass extracts such as those extracted from wood for this very reason. There tends to be a great amount of adsorption onto the column as the molar mass increases. Silica is the common choice for separation of tannins by size; however, other workers have used other stationary phases such as Kelm et al. who used a diol stationary phase to achieve the same type

of separation for a cocoa extract. The separation was achieved making use of a linear gradient with acetonitrile-aqueous methanol-acetic acid.

The elution order in reverse phase separation of condensed tannins is not related to the molar mass and can only be achieved up to tetramers, beyond this the molecules are too complex to achieve separation [26]. The isomers present in the procyanidins with the same molar mass can be resolved using reversed phase HPLC but molecules with different degrees of polymerization may overlap due to the fact that the number of isomers increases with an increase in DP. Guyot et al. separated a apple skin extract by reversed phase HPLC on a C18 column with acetic acid-methanol-water as mobile phase, separation was good for the lower molar mass oligomers but the polymeric fraction was not resolved [26] . Peng et al. analysed a grape seed extract and the same result was obtained, the polymeric fraction was not resolved in the chromatogram [30]. Thus it can be concluded that reversed phase LC is an excellent tool to analyse tannins by chemical composition but due to the complexity of these molecules it fails to separate the polymeric fraction contained in tannin extracts.

In order to identify the peaks in HPLC chromatograms, fractions are normally collected and analysed off-line by spectroscopic techniques such as NMR, MALDI-TOF and ESI-MS [7,52,65,98,123]. The ESI-MS analyses are often carried out on-line [23,26,33,42,49,57,116]. In the case of hydrolysable tannins, ^{1}H and ^{13}C NMR analyses are able to provide detailed structural assignments [95-97]. In order for this to be successful, the pure fraction of a specific structure is required. Correlation spectra such as COSY (Correlation Spectroscopy) and HMBC (Heteronuclear Multiple Bond Coherence) can be used to elucidate complex structures such as n-butylgallate 4-O-(2', 6'-di-O-galloyl)-β-D-glucopyranose [98] (**Figure 2.14**).

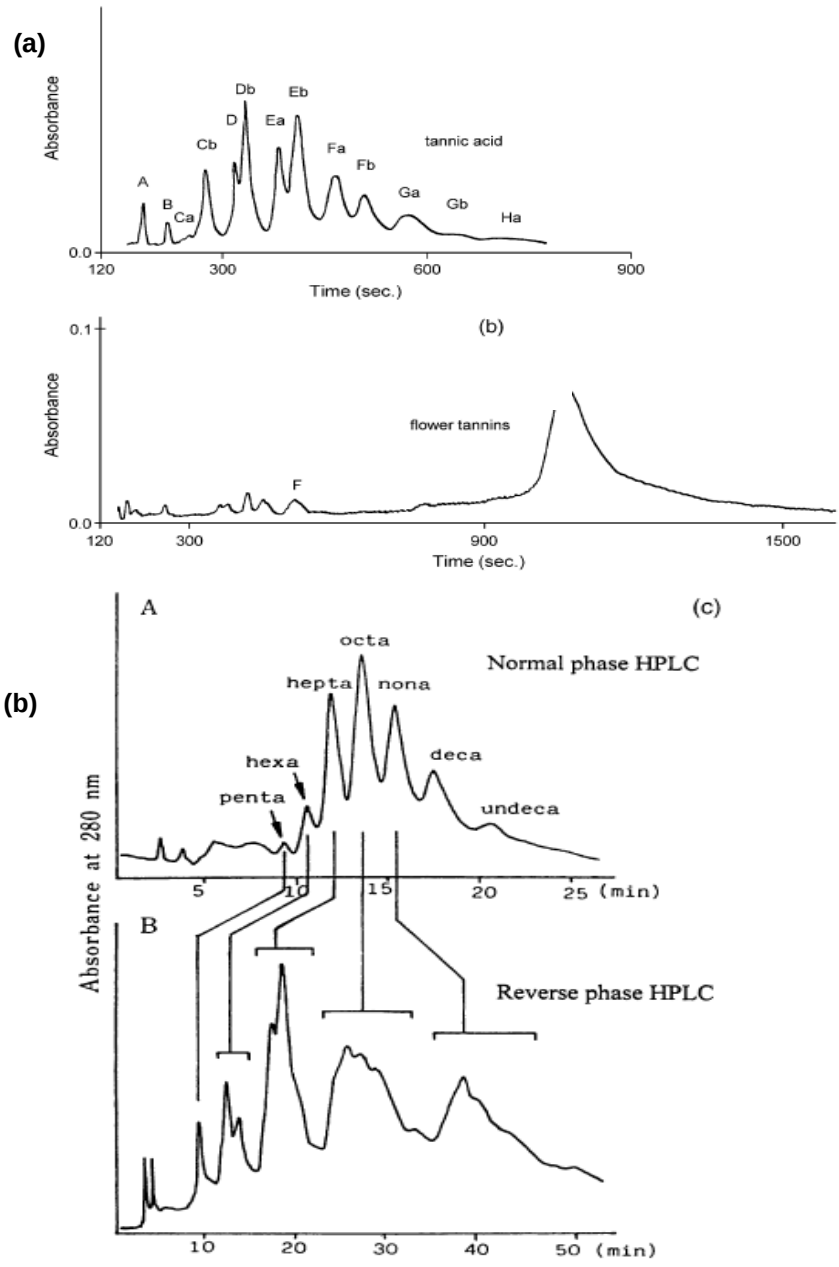


Figure 2.13: Normal phase HPLC of hydrolysable tannins from (a) tannic acid and fireweed flower tannins [58]. The letters were used to indicate peaks with the same retention time and subscripts describe a shoulder on the major peak, the tannic acid was used as a reference and the peaks appearing eluting at the same point in the flower tannin were assigned the same structure (b) Normal and reversed phase HPLC of galloylglucoses in leaves of *M. Indica* [120]. The penta, hexa, hepta etc, in the figure indicate the oligomer size, according to gallic acid residue chain length.

In the analysis of oligomeric tannins spectroscopic techniques are essential tools in the determination elution order and chemical structures of the molecules. Although in the case of higher molar mass tannin extracts HPLC gives incomplete separation, analysis by this method is still essential in understanding the chemical composition and molar mass distributions of these complex molecules.

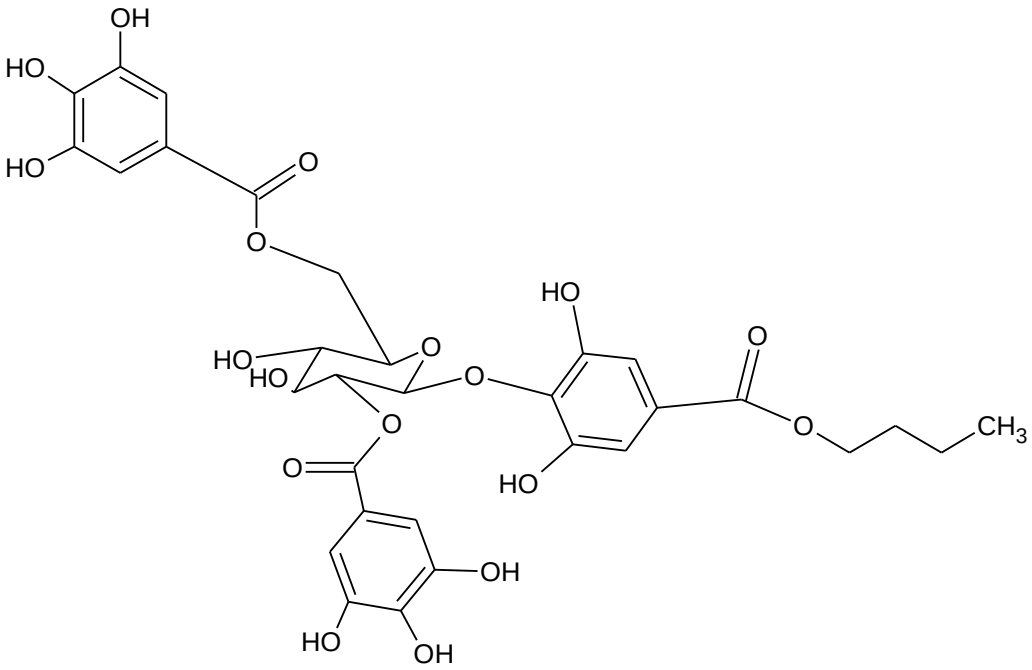


Figure 2.14: Chemical structure of *n*-butylgallate of 4-O-(2',6'-di-O-galloyl)-β-D-glucopyranose (hydrolysable tannin) elucidated by ^{13}C and ^{1}H NMR [98]

2.5. References

[1] A.E. Hagerman, in A.E. Hagerman (Editor), Tannin Handbook, http://www.users.muohio.edu/hagermae/tannin.pdf, Oxford, 1998.

[2] C.W. Oo, A. Pizzi, H. Pasch, M.J. Kassim, J Appl Polym Sci 109 (2008) 963.

[3] A. Pizzi, Int. J. Adhes. Adhes. 1 (1980) 13.

[4] E. Haslam, Practical polyphenolics from structure to molecular recognition and physiological action, Cambridge University Press, New York, 1998.

[5] G.E. Rohr, G. Riggio, B. Meier, O. Sticher, Phytochem. Anal. 11 (2000) 113.

[6] L.L. Zhang, Y.M. Lin, Molecules 13 (2008) 2986.

[7] I. Mueller-Harvey, Anim. Nutr. Feed Technol. 91 (2001) 3.

[8] P. Schofield, D.M. Mbugua, A.N. Pell, Anim. Feed Sci. Technol. 91 (2001) 21.

[9] A. Pizzi, Wood Adhesives Chemistry and Technology, Marcel Dekker, New York, 1983.

[10] J.-P. Salminen, T. Roslin, M. Karonen, J. Sinkkonen, K. Pihlaja, P. Pulkkinen, J. Chem. Ecol. 30 (2004) 1693.

[11] M. Noferi, E. Masson, A. Merlin, A. Pizzi, X. Deglise, J Appl Polym Sci 63 (1997) 475.

[12] A.E. Hagerman, K.M. Riedl, G.A. Jones, K.N. Sovik, N.T. Ritchard, P.W. Hartzfeld, T.L. Riechel, J. Agric. Food. Chem. 46 (1998) 1887.

[13] Z. Peng, Y. Hayasaka, P.G. Iland, M. Sefton, P. Hoj, E.J. Waters, J. Agric. Food Chem. 49 (2001) 26.

[14] K.J. Anderson, S.S. Teuber, A. Gobeille, P. Cremin, A.L. Waterhouse, F.M. Steinberg, J. Nutr. 131 (2001) 2837.

[15] I. Fecka, Phytochem. Anal. 20 (2009) 177.

[16] A. Pizzi, Advanced Wood Adhesives Technology CRC Press, New York, 1994.

[17] T. White, J. Sci. Food Agric. 8 (1957) 377.

[18] P.W. Hartzfeld, R. Forkner, M.D. Hunter, A.E. Hagerman, J. Agric. Food. Chem. 50 (2002) 1785.

[19] A. Romani, F. Ieri, B. Turchetti, N. Mulinacci, F.F. Vincieri, P. Buzzini, J. Pharm. Biomed. Anal. 41 (2006) 415.

[20] I.V. Popov, I.N. Andreeva, M.V. Gavrilin, Pharm. Chem. J. 37 (2003) 360.

[21] G. Piperopoulos, R. Lotz, A. Wixforth, T. Schmierer, K.-P. Zeller, J. Chromatogr. B: Biomedical Sciences and Applications 695 (1997) 309.

[22] C. Santos-Buelga, A. Scalbert, J. Agric. Food Chem. 80 (2000) 1094.

[23] J.M. Ricardo da Silva, J. Rigaud, V. Cheynier, A. Chemina, M. Moutounet, Phytochemistry 30 (1991) 1259.

[24] A. Scalbert, Phytochemistry 30 (1991) 3875.

[25] J.-L. Puech, F. Feuillat, J.R. Mosedale, Am. J. Enol. Vitic. 50 (1999) 469.

[26] V. Cheynier, J.-M. Souquet, E. Le Roux, S. Guyot, J. Rigaud, Methods Enzymol 299 (1999) 178.

[27] H.A. Stafford, H.H. Lester, Plant Physiol. 66 (1980) 1085.

[28] J.A. Kennedy, A.W. Taylor, J Chromatogr A 995 (2003) 99.

[29] L.Y. Foo, L.J. Porter, J. Sci. Food Agric. 32 (1981) 711.

[30] Z. Peng, Y. Hayasaka, P.G. Iland, M. Sefton, P. HÃ¸j, E.J. Waters, J. Agric. Food Chem. 49 (2000) 26.

[31] A. Yanagida, T. Kanda, T. Takahashi, A. Kamimura, T. Hamazono, S. Honda, J. Chromatogr. A 890 (2000) 251.

[32] S.A. Lazarus, G.E. Adamson, J.F. Hammerstone, H.H. Schmitz, J Agric Food Chem 47 (1999) 3693.

[33] C. Prieur, J. Rigaud, V. Cheynier, M. Moutounet, Phytochemistry 36 (1994) 781.

[34] A. Yanagida, T. Shoji, Y. Shibusawa, J. Biochem. Bioph. Methods 56 (2003) 311.

[35] H. Pasch, A. Pizzi, J. Appl. Polym. Sci. 85 (2002) 429.

[36] A. Pizzi, H. Pasch, K. Rode, S. Giovando, J. Appl. Polym. Sci. 113 (2009) 3847.

[37] A.E. Hagerman, M.E. Rice, N.T. Ritchard, J. Agric. Food. Chem. 46 (1998) 2590.

[38] H. Pasch, A. Pizzi, K. Rode, Polymer 42 (2001) 7531.

[39] J.F. Hammerstone, S.A. Lazarus, A.E. Mitchell, R. Rucker, H.H. Schmitz, J Agric Food Chem 47 (1999) 490.

[40] B. Labarbe, V. Cheynier, F. Brossaud, J.-M. Souquet, M. Moutounet, J. Agric. Food Chem. 47 (1999) 2719.

[41] C.G. Krueger, N.C. Dopke, P.M. Treichel, J. Folts, J.D. Reed, J. Agric. Food Chem. 48 (2000) 1663.

[42] M.A. Kelm, J.C. Johnson, R.J. Robbins, J.F. Hammerstone, H.H. Schmitz, J. Agric. Food. Chem. 54 (2006) 1571.

[43] N. Nuengchamnong, K. Ingkaninan, LWT Food Sci. Technol. 42 (2009) 297.

[44] P. Navarrete, A. Pizzi, H. Pasch, K. Rode, L. Delmotte, Ind. Crops Prod. 32 (2010) 105.

[45] N. Meikleham, A. Pizzi, A. Stephanou, J. Appl. Polym. Sci. 54 (1994) 1827.

[46] C. Pena, K. de la Caba, A. Retegi, C. Ocando, J. Labidi, J.M. Echeverria, I. Mondragon, J. Therm. Anal. Calorim. 96 (2009) 515.

[47] C.S. Ku, S.P. Mun, Wood Sci. Technol. 41 (2007) 235.

[48] K. Li, X. Geng, J. Simonsen, J. Karchesy, Int J Adhes Adhes 24 (2003) 327.

[49] K.M. Kalili, A. de Villiers, J. Chromatogr. A 1216 (2009) 6274.

[50] A. Yanagida, T. Shoji, T. Kanda, Biosci. Biotechnol., Biochem. 66 (2002) 1972.

[51] N. Vivas, M.-F. Nonier, N. Vivas de Gaulejac, C. Absalon, A. Bertrand, M. Mirabel, Anal. Chim. Acta 513 (2004) 247–256.

[52] B. Zywicki, T. Reemtsma, M. Jekel, J. Chromatogr. A 970 (2002) 191.

[53] A. Pizzi, D. Thompson, J. Appl. Polym. Sci. 55 (1995) 107.

[54] M. Theis, B. Grohe, Holz als Roh- und Werkstoff 60 (2002) 291.

[55] F. Pichelin, C. Kamoun, A. Pizzi, Holz als Roh- und Werkstoff. 57 (1999) 305.

[56] U. Svedstrom, H. Vuorela, R. Kostiainen, K. Huovinen, I. Laakso, R. Hiltunen, J Chromatogr A 968 (2002).

[57] J. Rigaud, M.T. Escribano-Bailon, J. Prieur, J.-M. Souquet, V. Cheynier, J Chromatogr A 654 (1993) 255.

[58] A.E. Hagerman, C.T. Robbins, Y. Weerasuriya, T.C. Wilson, C. McArthur, J. Range Manage. 45 (1992) 57.

[59] T. Hatano, M. Hori, R.W. Hemingway, T. Yoshida, Phytochemistry 63 (2003) 817.

[60] J.-P. Salminen , V. Ossipov, J. Loponen, E. Haukioja, K. Pihlaja, J. Chromatogr. A 864 (1999) 283.

[61] I.J.A. Baker, B. Matthews, H. Suares, I. Krodkiewska, D.N. Furlong, F. Grieserb, C.J. Drummond, Journal of Surfactants and Detergents 3 (2000) 1.

[62] L.J. Porter, L.Y. Foo, Phytochemistry 21 (1980) 2947.

[63] V. Nunez, C. Gomez-Cordoves, B. Bartolome, Y.-J. Hong, A.E. Mitchell, J. Sci. Food Agric. 86 (2006) 915.

[64] M. Markom, M. Hasan, W.R.W. Daud, H. Singh, J.M. Jahim, Sep. Purif. Technol. 52 (2007) 487.

[65] C. Perret, R. Pezet, R. Tabacchi, Phytochem. Anal. 14 (2003) 202.

[66] B. Sun, C. Leandro, J.M. Ricardo da Silva, I. Spranger, J. Agric. Food Chem. 46 (1998) 1390.

[67] S. Taniguchi, Y. Imayoshi, R. Yabu-uchi, H. Ito, T. Hatano, T. Yoshida, Phytochemistry 59 (2002) 191.

[68] A. de Villiers, F. Lynen, A. Crouch, P. Sandra, Chromatographia 59 (2004) 403.

[69] A. Yanagida, T. Kanda, T. Shoji, M. Ohnishi-Kameyama, T. Nagata, J. Chromatogr. A 855 (1999) 181.

[70] A. Pizzi, Colloid. Polym. Sci. 257 (1979) 37.

[71] A. Pizzi, Int. J. Adhes. Adhes. (1981) 213.

[72] S. Giovando, A. Pizzi, H. Pasch, K. Rode, J. Appl. Polym. Sci. 114 (2009) 1339.

[73] A. Behrens, N. Maie, H. Knicker, I. Ko¨ gel-Knabner, Phytochemistry 62 (2003) 1159.

[74] Y. Brun, P. Alden, J. Chromatogr. A 966 (2002) 25.

[75] G. Guiochon, J. Chromatogr. A 1126 (2006) 6.

[76] T. Chang, J. Polym. Sci. Part B: Polym. Phys 43 (2005) 1591–1607.

[77] H. Pasch, B. Trathnigg, HPLC of Polymers, New York : Springer, Berlin, 1998.

[78] H.J.A. Philipsen, J Chromatogr A 1037 (2004) 329.

[79] Y. Brun, J. Liq. Chromatogr. Related Technol. 22 (1999) 3067.

[80] Y. Brun, J. Liq. Chromatogr. Related Technol. 22 (1999) 3027.

[81] J.-A. Raust, A. Brüll, C. Moireb, C. Farcet, H. Pasch, J Chromatogr A 1203 (2008) 207–216.

[82] P. Hemstrom, K. Irgum, J. Sep. Sci. 29 (2006) 1784.

[83] A. Yanagida, H. Murao, M. Ohnishi-Kameyama, Y. Yamakawa, A. Shoji, M. Tagashira, T. Kanda, H. Shindo, Y. Shibusawa, J. Chromatogr. A 1143 (2007) 153.

[84] A.J. Alpert, J. Chromatogr. 499 (1990) 177.

[85] T. Ikegami, H. Fujita, K. Horie, K. Hosoya, N. Tanaka, Anal. Bioanal. Chem. 386 (2006) 578.

[86] H. Lindner, B. Sarg, C. Meraner, W. Helliger, J. Chromatogr. A 743 (1996) 137.

[87] P. Ciminiello, C. Dell'Aversano, E. Fattorusso, M. Forino, G.S. Magno, L. Tartaglione, M.A. Quilliam, A. Tubaro, R. Poletti, Rapid Commun. Mass Spectrom. 19 (2005) 2030.

[88] M. Lazzari, G. Liu, S. Lecommandoux, Block Copolymers in Nanoscience, Wiley-VCH, 2006

[89] T. Ikegami, K. Tomomatsu, H. Takubo, K. Horie, N. Tanaka, J. Chromatogr. A 1184 (2008) 474.

[90] D.V. McCalley, U.D. Neueb, J. Chromatogr. A 1192 (2008) 225.

[91] V.V. Tolstikov, O. Fiehn, Anal. Biochem. 301 (2002) 298.

[92] M.-K. Lai, J.-Y. Wang, R.C.-C. Tsiang, Polymer 46 (2005) 2558.

[93] J.L. Wolfender, S. Rodriguez, K. Hostettmann, J. Chromatogr. A 794 (1998) 299.

[94] T. Tanaka, T. Sueyasu, G.-I. Nonaka, I. Nishioka, Chem. Pharm. Bull. 32 (1984) 2676.

[95] J.L. Wolfender, K. Ndjoko, K. Hostettmann, Phytochem. Anal. 12 (2001) 2.

[96] M. Sandvoss, L.H. Pham, K. Levsen, A. Preiss, C. Mügge, G. Wünsch, ChemInform 2000 (2000) 1253.

[97] E. de Rijke, P. Out, W.M.A. Niessen, F. Ariese, C. Gooijer, U.A.T. Brinkman, J. Chromatogr. A 1112 (2006) 31.

[98] M.S.A. Marzouk, Phytochem. Anal. 19 (2008) 541.

[99] L. Kong, X. Li, H. Zou, H. Wang, X. Mao, Q. Zhang, J. Ni, J. Chromatogr. A 936 (2001) 111.

[100] W. Li, J.F. Fitzloff, J. Liq. Chromatogr. Related Technol. 25 (2002) 2501.

[101] R. Lucena, S. Cardenas, M. Gallego, M. Valcarcel, J. Chromatogr. A 1081 (2005) 127.

[102] J. Lokvam, T.A. Kursar, J. Chem. Ecol. 31 (2005) 2563.

[103] K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida, T. Matsuo, Rapid Commun. Mass Spectrom. 2 (1988) 151.

[104] M. Karas, U. Bachmann, U. Bahr, F. Hillenkamp, Int J Mass Spectrom Ion Processes 78 (1987) 53.

[105] F. Hillenkamp, M. Karas, R.C. Beavis, B.T. Chait, Anal Chem 63 (1991) 1193A.

[106] O.-K. Mayumi, Y. Akio, K. Tomomasa, N. Tadahiro, in, 1997, p. 31.

[107] G. Montaudo, M.S. Montaudo, S. Samperi, Mass Spectrometry of Polymers, CRC Press, Boca Raton, 2002.

[108] J. Schweer, J. Sarnecki, F. Mayer-Posner, K. Mullen, H. Rader, J. Spickermann, Macromolecules 29 (1996) 4536.

[109] P. Danis, D. Karr, Macromolecules 28 (1995) 8548.

[110] S.D. Hanton, Chem. Rev. 101 (2001) 527.

[111] B. Thomson, K. Suddaby, A. Rudin, G. Lajoie, Eur Polym J 32 (1996) 239.

[112] O. Belgacem, A. Bowdler, I. Brookhouse, F.L. Brancia, E. Raptakis, Rapid Commun. Mass Spectrom. 20 (2006) 1653.

[113] A.I.T. Jackson, K.R. Jennings, J.H. Scrivens, J am Soc Mass Spectrom 8 (1997) 76.

[114] R.J. Cotter, W. Griffith, C. Jelinek, J. Chromatogr. B 855 (2007) 2.

[115] R. Saf, C. Mirtl, K. Hurnmel, Acta Polym. 48 (1997) 513.

[116] R. Medrano, M.T.R. Laguna, E. Saiz, M.P. Tarazona, Phys Chem Chem Phys 5 (2003) 151.

[117] R.S. Thompson, D. Jacques, E. Haslam, R.J.N. Tanner, J. Chem. Soc., Perkin Trans. 1 (1972).

[118] J. Rigaud, J. Perez-Ilzarbe, J.M. Ricardo da Silva, V. Cheynier, J Chromatogr A 540 (1991) 401.

[119] T.C. Wilson, A.E. Hagerman, J. Agric. Food Chem. 38 (1990) 1678.

[120] T. Okuda, T. Yoshida, T. Hatano, J. Nat. Prod. 52 (1989) 1.

[121] Z. Czochanska, L.Y. Foo, R.H. Newman, L.J. Porter, J. Chem. Soc., Perkin Trans 1 (1980) 2278.

[122] V. Cheynier, M. Duenas-Paton, E. Salas, C. Maury, J.-M. Souquet, P. Sarni-Manchado, H. Fulcrand, Am. J. Enol. Vitic. 57 (2005) 298.

[123] M.A.M. Nawwar, M.S. Marzouk, W. Nigge, M. Linscheid, J. Mass Spectrom. 32 (1997) 645.

Chapter 3

Structural Elucidation of Polyflavonoids by Spectroscopic Techniques

3. 1. Introduction

Polyflavonoid tannins are natural polyphenolic materials that are widely distributed in the plant kingdom as secondary metabolites [1]. The extracts obtained from wood are used industrially in order to synthesise tannin-based wood adhesives either by auto-condensation or reaction with formaldehyde [2-5]. In the synthesis of wood adhesives some various extraction modes may be used and these may alter the structure of the tannin molecules. These synthesis intermediates have been shown to increase reactivity with formaldehyde, thus understanding the specific structure is significant. As a result considerable amount of research has gone into the study of the structure-property relationship of these tannins. The repeat unit which is a flavan-3-ol is shown in **Figure 3.1**. In this structure, 2 phenolic rings are present, namely the A and B-rings. Polyflavonoid oligomers and polymers are made up of these monomers which are linked through C4-C8 or C4-C6 bonds, with the former being most common. These types of oligomers are known as B-type oligomers; the A-type oligomers are less common and formed through the linkage C2-O7 of the flavonoid units [6].

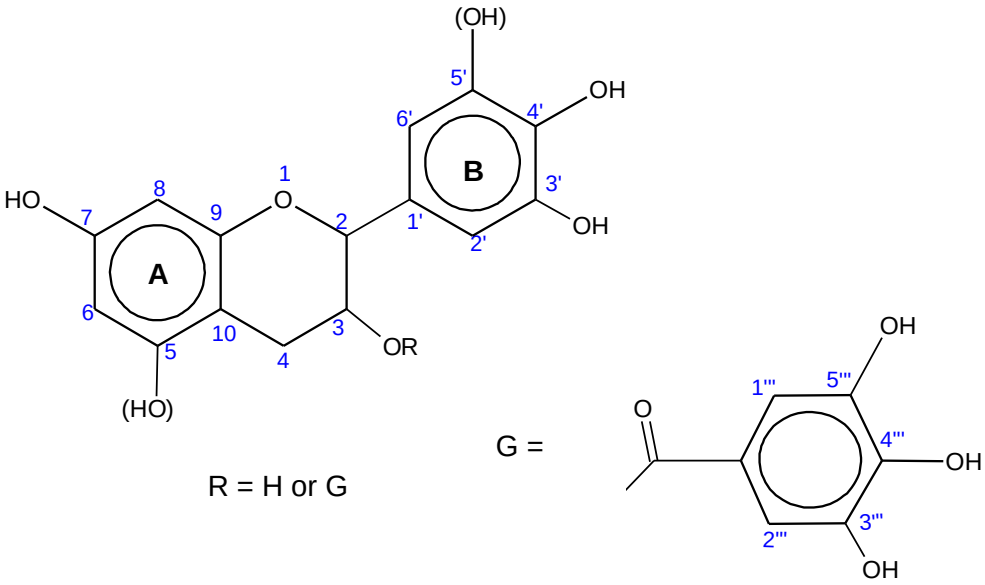


Figure 3.1: Structure of a flavan-3-ol monomer.

Tannin analysis has been a matter of interest due to the fact that they show varying structural properties depending on the type of plant and conditions of extraction used. This variation in chemical composition and molar mass distribution has been widely studied; however, very few strides have been made in terms of fully

describing the structural content of tannin extracts. In view of this, numerous analytical methods have been developed in order to analyse these complex molecules. The most common methods make use of spectroscopic techniques for determining the average chemical composition of tannin extracts. ^{1}H and ^{13}C NMR spectrometry have been the most commonly used techniques. Since its introduction MALDI-TOF MS by Tanaka as well as Karas and Hillenkamp, the use of this techniques has widely spread [7-16]. The use of ^{13}C NMR is limited due to the fact that broad peaks are obtained in the case of high molar mass tannin extracts, this makes interpretation difficult. Although this is the case, this technique can provide valuable average information in terms of the isomers present and molar mass. ESI-MS is also widely used in tannin analysis; in this case the use is limited by the formation of multiply charged ions even in the negative mode, which is meant to be the simpler ion formation mechanism in the case of proanthocyanidins. In the positive mode some salt adducts may form and thus some methods need to be employed to suppress formation of these adducts. The presence of potassium and sodium adducts may further complicate the interpretation of the spectrum. For tannin analysis, specifically, MALDI-TOF has proven to be superior since it can provide information on structure and molar mass in a single experiment. This technique is not limited by complexity as are other averaging techniques. However, its one limitation is that it is unable to distinguish between structures that have the same molar mass but differing chemical compositions.

The combination of multiple spectroscopic techniques such as ^{13}C NMR, MALDI-TOF MS and ESI-MS can provide conclusive information on molar mass and chemical composition distributions [16]. Therefore this information can be useful in elucidating structures of unknown samples. The aim of this study was to show the application of these techniques in order to determine the chemical composition and molar masses of different types of commercial condensed tannins.

3. 2. Experimental

3.2.1. Reagents and materials

Cacao beans were purchased from a local supermarket. The (+)-epicatechin hydrate standard and HPLC grade acetone were purchased from Sigma-Aldrich (Steinheim, Germany) Deuterated acetone was also purchased from the same supplier. Deionised water was obtained using a Milli-Q water purification system (Millipore, Milford, MA, USA). The solutions prepared for NMR analysis were filtered through a 0.45 µm filters (Millipore). Due to the high concentrations used for NMR analysis an ultrasonic bath (Integral systems) was used in order to ensure maximum dissolution of the samples.

3.2.2. Instrumentation

3.2.2.1. ^{13}C NMR analyses

The ^{13}C NMR (300 or 400 MHz) analyses were recorded on a Varian 300 VNMRS instrument and Varian Unity Inova 400, respectively. Trimethylsiloxane (TMS) was used as the internal standard. The samples were run over the weekend whereas the standards were run overnight. The decoupling mode with nuclear overhauser effect (NOE) was used.

3.2.2.2. MALDI-TOF MS analyses

MALDI-TOF analyses were performed on a AximaTOF^2 spectrometer (Shimadzu Biotech, Manchester, UK), equipped with a nitrogen laser (337 nm), the pulsed extraction ion source accelerated the ions to a kinetic energy of 20 keV. All analyses were carried out in the linear positive mode. Calibration was done using 1450 Da polyethylene glycol (PEG) standard. The instrument is equipped with an ion gate in order to select precursor ions and a collision cell. Argon was used as the collision gas and the precursor ions were separated by approximately 4 Da mass window. The pressure in the collision cell was 8×10^{-6} mbar.

3.2.2.2. Direct injection ESI-MS analyses

The samples were introduced for analysis on the Waters Ultima API quadrupole time-of-flight (Q-TOF) mass spectrometer by direct injection through Waters Acquity UPLC system equipped with a binary solvent manager at 300 µL/min. The capillary voltage was kept at 3.5 kV and cone voltage of 35 and desolvation temperature was 250°C. Analyses were carried out in the positive and negative modes. The mass range that was scanned was 100-2000 amu. Data acquisition and processing was carried out on the MassLynx v4.0 software (Waters). Calibration was done using NaF. N_2 was used as the gas and for desolvation it was kept at 250 L/h and the cone was 50 L/h. The source temperature was 80°C.

3.2.3. Sample preparation in spectroscopic techniques

3.2.3.1. ^{13}C NMR analyses

A solution of 50/50 (% v/v) acetone/water was prepared, followed by preparation of a mixture of d_6-acetone/D_2O in the same ratio. These two solutions were combined and 1 mL of this mixture was used to dissolve 300 mg of each sample. In order to facilitate dissolution the samples were placed in an ultrasonic bath for 5 minutes. This allowed for further dissolution; however at this relatively high concentration the tannins were only partially soluble. The resulting sample solutions were the filtered using 0.45 µm nylon filter then transferred to an NMR tube to be analysed. The standards were prepared in pure d_6-acetone solution and were not subjected to any sonication since they are low molar mass compound thus making them readily soluble.

3.2.3.2. MALDI-TOF MS analyses

The samples were dissolved in acetone/water 50/50 (% v/v) at a concentration of 4 mg/mL. The sample solutions were mixed with a 10 mg/mL solution of the matrix in the same solvent. 2,5-dihydroxy benzoic acid

was used as the matrix. NaCl was added as the salt to enhance ion formation. The sample and matrix were then combined at ratios of 1:1. 5 µL of the resulting solution was spotted on a 100 well MALDI-TOF plate, followed by evaporation of the solvent at ambient temperature without any assistance. The MALDI-TOF target was then analysed to give the resulting spectra.

3.2.3.2 Direct ESI-MS analyses

For direct inject ESI-MS the samples were dissolved in a solution of mobile phase that consisted of (A) acetonitrile and acetic acid (99:1, %v/v) and (B) methanol, water and acetic acid (94.05: 4.95:1, % v/v/v). 25 % A and 75% B were actually used for sample dissolution in order to resemble the chromatographic system. The sample was prepared at a concentration of 10 mg/mL and only 2 µL of this solution was injected for analysis.

3. 3. Results and discussion

3.3.1. Analysis of polyflavonoids by ^{13}C NMR

^{13}C NMR spectroscopy is used to examine the chemical composition of various polyflavonoid extracts. As shown in **Figure 3.2** the structure of the tannins is more complex than that of the catechin monomer unit. In addition, the peaks observed in the extract show a variation from the simple procyanidin pattern that would be formed by a catechin monomer. The samples that will be considered in this discussion are water-extracted mimosa tannin, a cacao extract, and a maleic anhydride/NaOH modified quebracho extract.

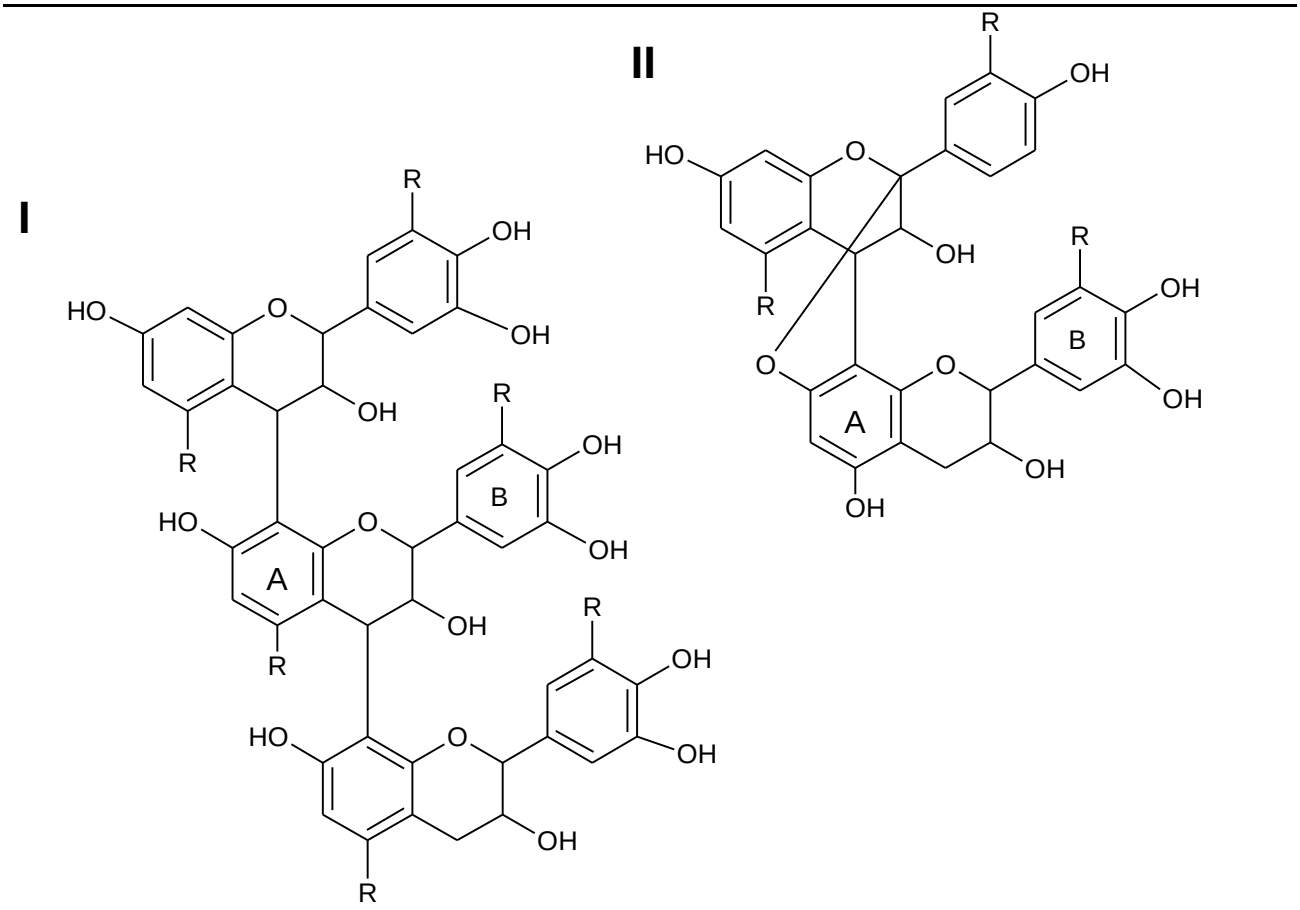


Figure 3.2: Structure of a (I) C4-C8 linked trimer [B-type] and (II) C2-O7 linked dimer [A-type]. R can represent an OH or H group.

The variations in composition of the different extracts are shown by the presence or absence and intensities of the appropriate peaks. The differences in chemical composition of industrial quebracho and mimosa extracts have already been reported by Pizzi and Stephanou [17]. The cacao extract is not discussed in detail since a significant amount of information is already available on its structure [18-20]. Due to this fact this tannin was used as a starting point to develop the method and subsequently compared to the other extracts.

The cacao extract is composed mainly of procyanidins and some prodelphinidins. Procyanidins (PC) are composed of two –OH groups on the A- and B-rings, whereas the prodelphinidins (PD) contain an additional –OH group on the B-ring. A standard, catechin hydrate was used to compare the structural content of the cacao extract with the mimosa and quebracho extract. A detailed look at the ^{13}C NMR spectra reveals some similarity in composition; however some important differences are also highlighted.

Table 3.1.: Sample list with description and tannin type.

Sample	Description	Type of tannin
6	Solvent extracted chestnut tannin	Hydrolysable tannin
7	Solvent extracted tara tannin	Hydrolysable tannin
8	Solvent extracted chinese gall tannin	Hydrolysable tannin
9	Solvent extracted turkey gall tannin	Hydrolysable tannin
10	Water extracted oak tannin	Hydrolysable tannin
11	Solvent extracted oak tannin	Hydrolysable tannin
12	Solvent extracted quebracho tannin	Condensed tannin
13	Ethanol extracted tara tannin	Hydrolysable tannin
14	Water extracted quebracho tannin	Condensed tannin
15	Water extracted (bisulphited) quebracho tannin	Condensed tannin
16	Modified quebracho tannin (product of reaction with maleic anhydride and NaOH)	Condensed tannin
17	Modified quebracho tannin	Condensed tannin
18	Water extracted mimosa tannin (bisulphited)	Condensed tannin
19	Water extracted mimosa tannin	Condensed tannin
23	Catechin hydrate	Model compound
24	(-)-Epicatechin	Model compound
25	Cacao Extract	Condensed tannin

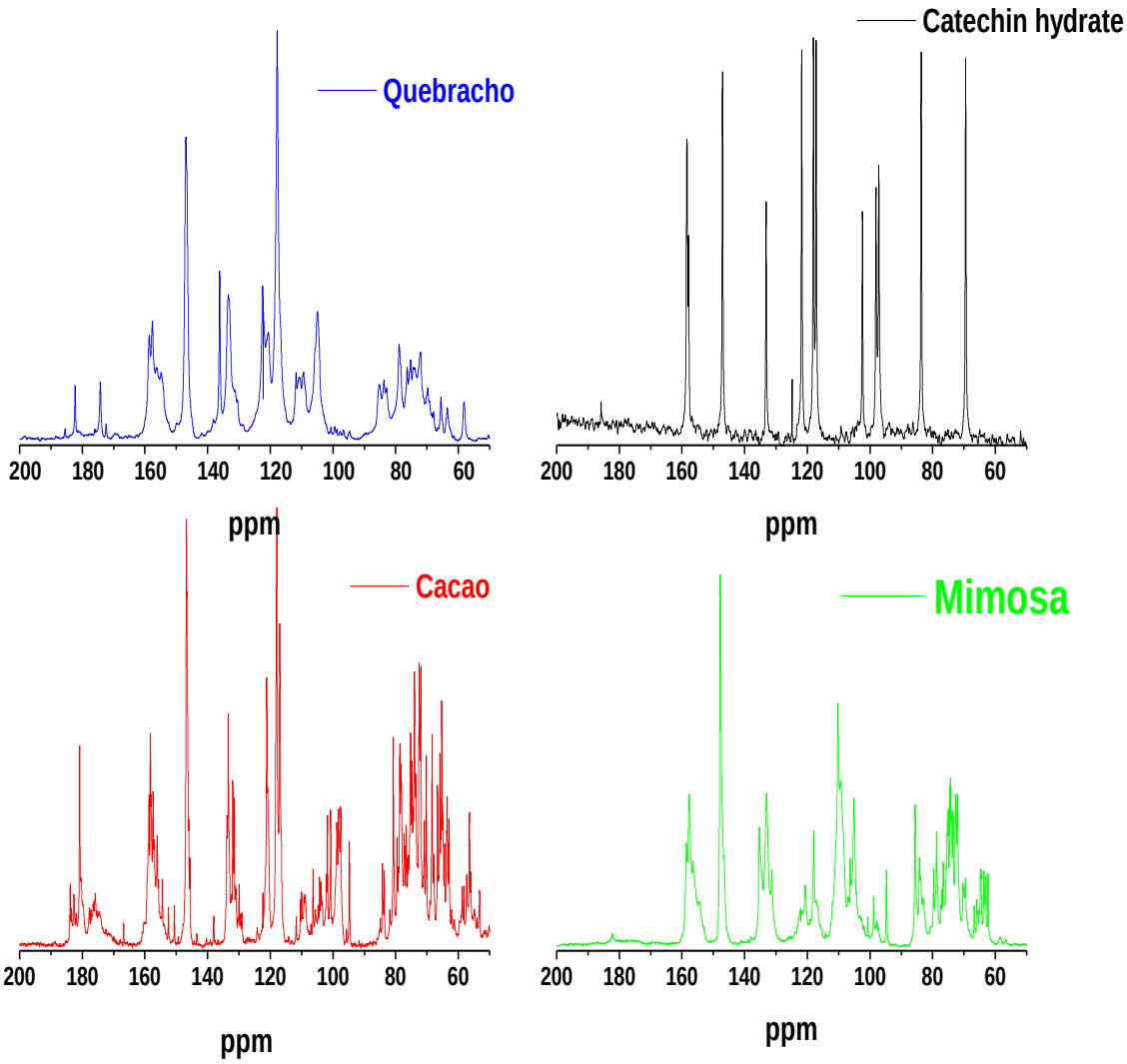


Figure 3.3: Liquid–state ^{13}C NMR spectra of condensed tannins extracted from quebracho (sample 16), mimosa (sample 19) and cacao (sample 25) with a catechin hydrate standard.

An overlay of the full spectra from the three samples as well as the catechin hydrate monomer is shown in **Figure 3.3**. The peaks that are of interest are the ones appearing in the region of 172-184 ppm which are present only in the cacao and quebracho extract (**Figure 3.4**). The presence of peaks in this region indicates two possibilities in structure, the first being the presence of a gallic acid residue at C3 position of the heterocycle and second they are an indication for quinone structures formed by the oxidation of the phenolic hydroxy groups that are shown in **Figure 3.1** [16,21]. The peak for catechin or epicatechin gallate appears at 178 ppm as a broad signal and in this region in the spectra shown in **Figure 3.4** the broad signal is split, this

could be due the different isomeric structures present for this carbon [16,21]. In quebracho these structures have been observed in low levels, their presence in this extract was unexpected. However, since these signals may overlap with the quinone structures formed by oxidation, it can be concluded that the presence of these peaks indicates extensive degradation that occurred in the extraction process (**Figure 3.4**). A catechin gallate monomer was analysed and the spectrum is similar to that shown for the catechin monomer, however an additional peak at 166 ppm belonging to the carbonyl of the gallic acid residue is observed. When the three tannins are compared to this standard in the respective region, it can clearly be seen that for the cacao tannin, the peak also appears at 166 ppm, however for the quebracho the signals are located more downfield.

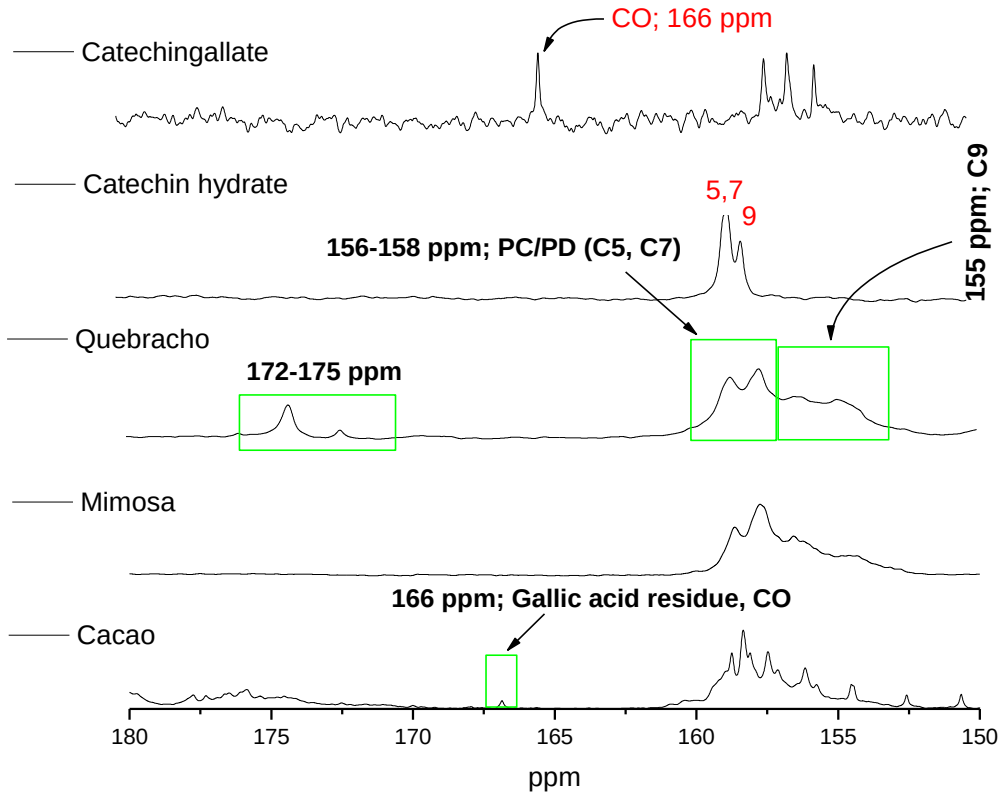


Figure 3.4: Liquid–state ^{13}C NMR spectra of condensed tannins extracted from quebracho (sample 16), mimosa (sample 19) and cacao tannin (sample 25) with a catechin hydrate and catechin gallate standards. Region: 180-150 ppm.

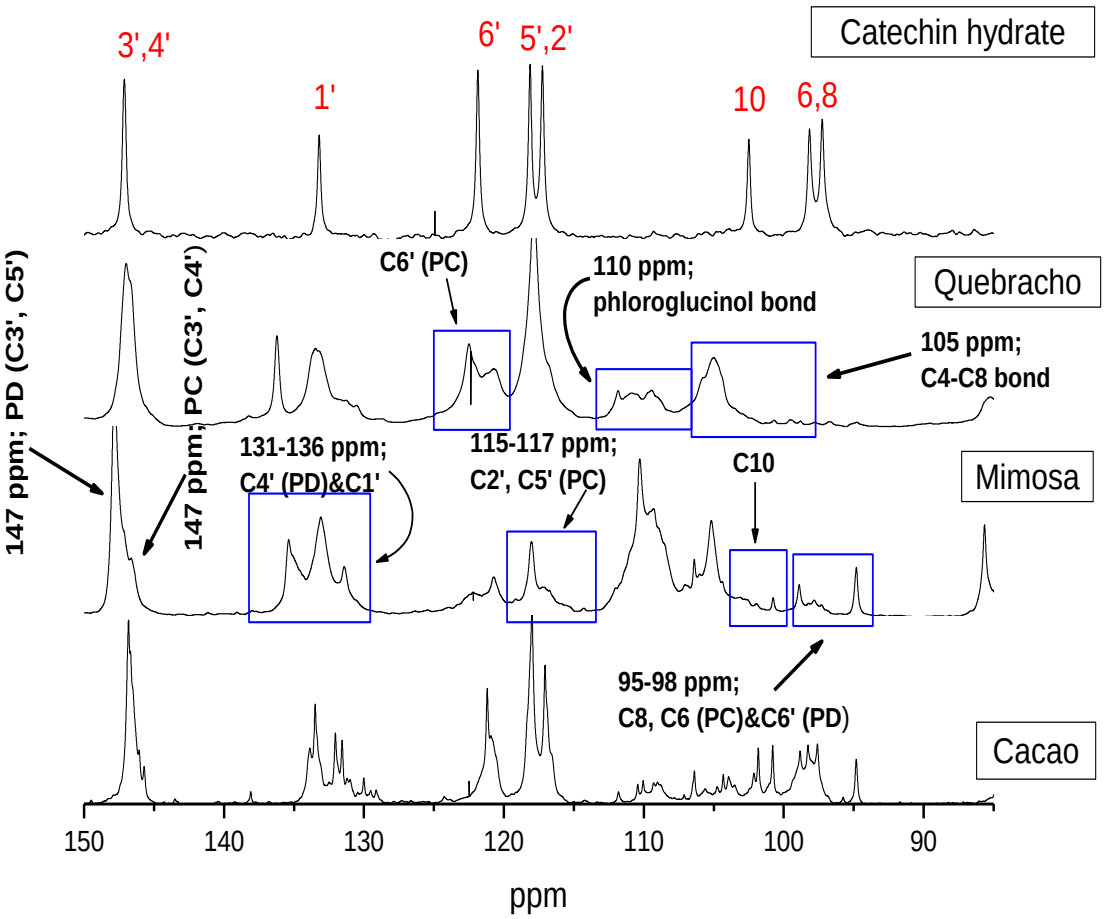


Figure 3.5 Liquid–state ^{13}C NMR spectra of condensed tannins extracted from quebracho (sample 16), mimosa (sample 19) and cacao tannin (sample 25) with a catechin hydrate standard. Region: 85-150 ppm.

The peak intensities of specific signals in ^{13}C NMR of tannin extracts can give an indication of the type of ring structures that are predominant in that structure, the relative degree of polymerisation and the type of bonds formed in the condensation reaction. The peaks indicating the free C6 and C8 carbons appear at 96-98 ppm and 95-96 ppm, respectively. The relative intensity of these peaks can give an indirect indication of the degree of polymerisation of a particular extract [22]. As seen in **Figure 3.5**, the mimosa and cacao extracts have relatively intense peaks in this region whereas the quebracho extract shows almost no signal. This then infers that quebracho is more polymerised than both cacao and mimosa extract, a similar trend to what was shown by Pizzi and Stephanou for quebracho and mimosa. Mimosa extract has the highest C8 (95-96 ppm) peak intensity which suggests that it forms mainly the C4-C6 bonds and cacao forms C4-C8 bonds since the

C6 peak intensity is higher in this extract, this observation was expected because procyanidins form mainly C4-C8 bonds [8].

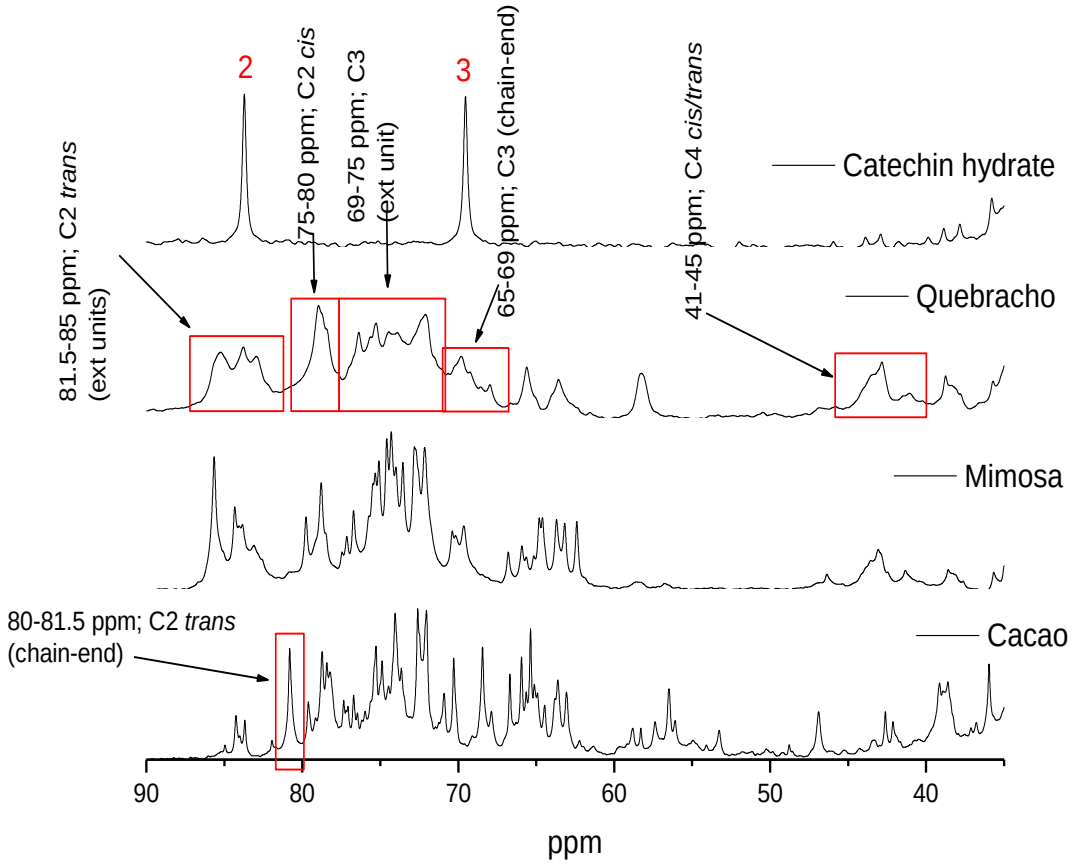
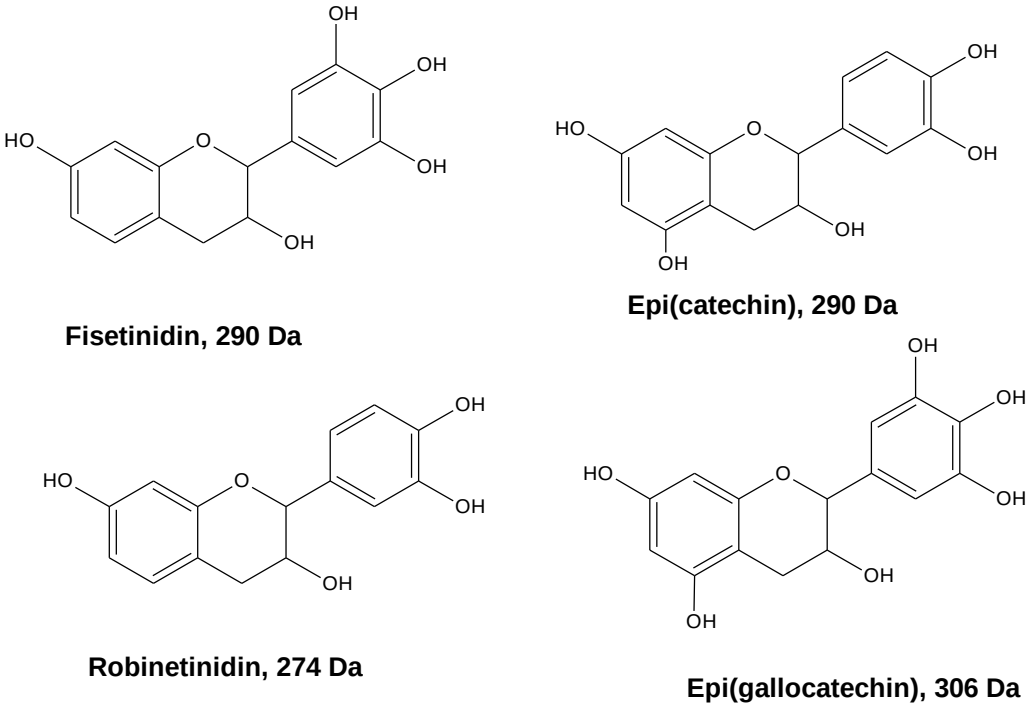


Figure 3.6: Liquid–state ^{13}C NMR spectra of condensed tannins extracted from quebracho (sample 16), mimosa (sample 19) and cacao tannin (sample 25) with a catechin hydrate standard. Region: 90-35 ppm.

The relative DP can also be shown by the peak appearing at 110 ppm which indicates the amount of phloroglucinol rings that form the interflavonoid bonds, shown in **Figure 3.5** [17,22]. The highest intensity of this peak is observed in the mimosa extract which implies that this tannin contains higher amounts of the phloroglucinol A-ring; therefore, it contains the lowest amount of procyanidin/prodelphinidin molecules. The B-ring structure may be established by observing the intensity of the peaks around 116 ppm which represent the hydroxylated C5' and C2' carbons in prodelphinidins. The highest intensity is observed in quebracho and

cacao extracts, this information observed in the respective regions, therefore suggests that the cacao and quebracho extracts have the highest proportion of prodelphinidin units. Another method of obtaining the same information is to integrate the peaks in the region of 145 ppm which represents the C3' and C4' carbons of the catechin structure.

The region between 35 and 90 ppm can give an indication of the isomeric structures that are present in a tannin extract (**Figure 3.6**). The peaks between 60 and 80 ppm, show how complex the samples are. In the case of the standard, there are only 2 peaks for the C2 and C3 of the catechin monomer. These two carbons can be either *cis* or *trans* and depending on the location the chemical shift is different.



Fisetinidin, 290 Da

Epi(catechin), 290 Da

Robinetinidin, 274 Da

Epi(gallocatechin), 306 Da

Figure 3.7: Structures of monomers present in the condensed tannin extracts with their molar mass indicated.

3.3.2. MALDI-TOF analysis on proanthocyanidins

Quebracho and mimosa tannins are extracted in large amounts industrially and the cacao extract has been extensively studied in the field of food science [18,23,24].Quebracho and mimosa tannins are known to be

composed of profisetinidin and prorobinetinidin oligomers. The cacao tannin consists of the simpler counter parts, procyanidin (Figure 3.8). The major peaks of the quebracho extract are separated by mass increment of 272 Da, belonging to a fisetinidin monomer (A-type structure) whereas mimosa and cacao have a major distribution separated by 288 Da which may be a catechin or prorobinetinidin monomer (B- type structure) (**Figure 3.7**).

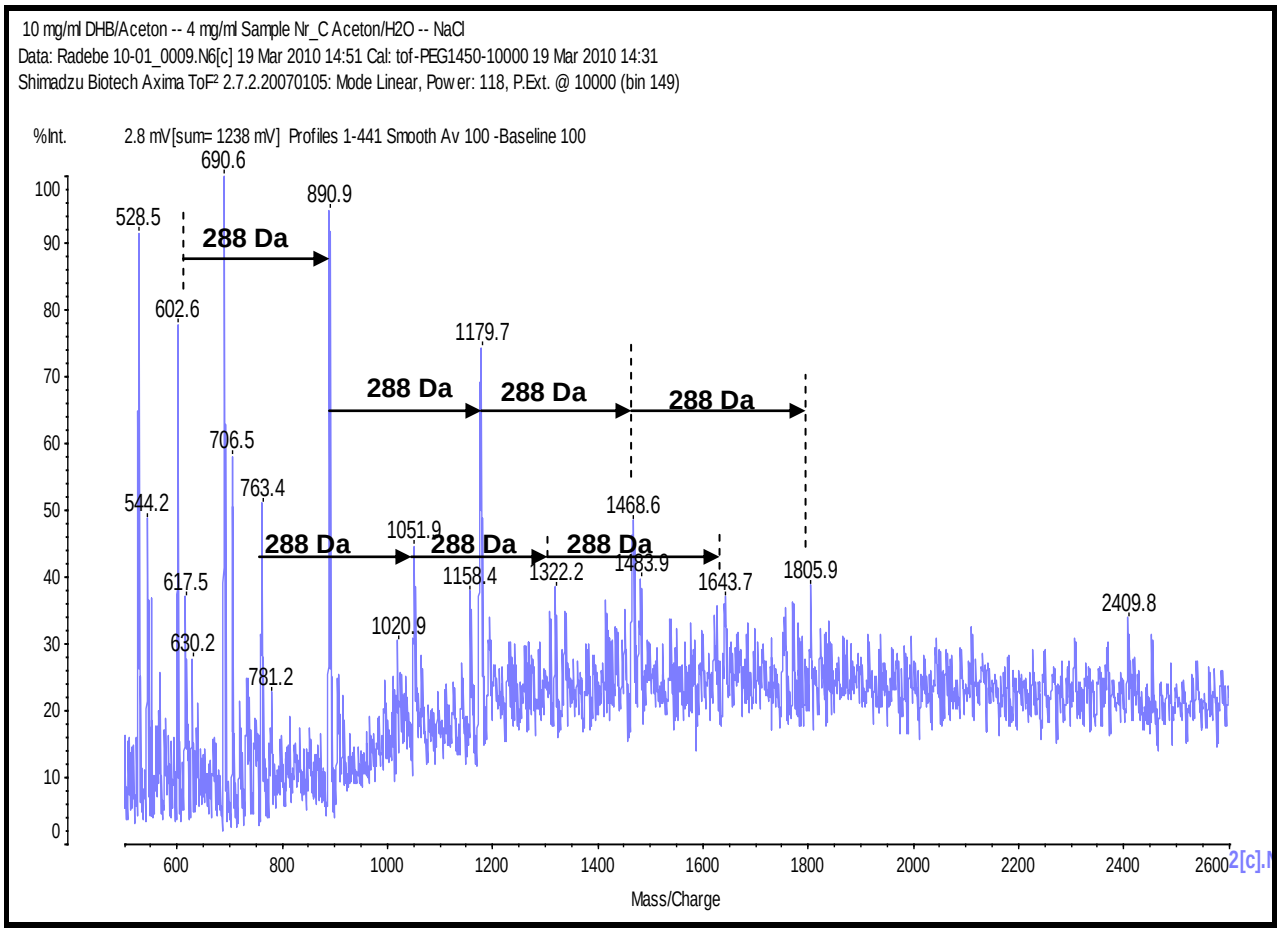


Figure 3.8: MALDI-TOF spectrum of the cacao tannin extract. Mass range: 500-2600 Da.

The cacao is known to be a procyanidin therefore the B-type structure in this case represents a catechin monomer. The calculated and observed masses following a MALDI-TOF analysis of the cacao extract is shown in **Table 3.2**. Another set of peaks is detected 152 Da from the major series and the oligomers therefore have the same structure but differ in end groups. An example would be the peak at 1052 Da which has a mass increment of 152 Da from the dominant trimer at 891 Da. The 152 Da belongs to a galloyl unit attached to the C3 carbon of the basic structure. The presence of these types of structures is well documented for other types of extracts [21,25]. Therefore in order to account for this difference the equation

used to calculate the theoretical masses must be adjusted and thus the following equation is used: $M+Na^+$ = 23 (Na) + 2 (2 x H endgroups) + 272A + 288B + 304C. The results of this calculation and peak assignment are shown in **Table 3.2**.

3.3.2.2. Quebracho extracts: polyflavonoid mass profiles of extracts obtained by various methods, MALDI-TOF results

For all the quebracho extracts (**Figure 3.9**) the main signals are separated by 272 Da, this mass corresponds to a profisetinidin monomer. The water extracted quebracho tannin (sample 14) is shown in **Figure 3.9a**; the spectrum shows an additional series which also has the same repeat unit as the main series. For example, 1132-860 Da and 1308-1036 Da, give the same value of 272 Da. This second minor series has the same structure, though there is a difference in end groups. This set of peaks represents oligomers that have undergone degradation reaction. This occurs via an elimination reaction favouring the removal of the A-ring and the C4 from the heterocycle of the proanthocyanidin structure. This explanation stems from the observation made by Pasch et al. that industrially extracted quebracho extracts contain degraded oligomers [12]. In addition, there are other significant peaks that are observed in the MALDI-TOF spectrum which have a 16 Da mass difference from the peaks forming the major series. These peaks indicate the presence of other types of structures such as robinetinidol, which consists of an additional hydroxyl group on the B-ring forming the B-type structures [12,25] . Although in this case the same mass of 288 Da given by this monomer can also be attributed to catechin, the quebracho tannin has been shown to consist of profisetinidins and probinetinidins [17]. In order to calculate the possible structure combinations the following equation, $M+Na^+$ = 23 (Na) + 2 (endgroups, 2 x H) + 272A + 288 B (**Table 3.3**). This calculation was used in order to predict the structures forming the major series in each extract. The C-type structure (prodelphinidins) was excluded from the calculation due to the fact that these structures are known not to be present in this extract [12,25]. Using this equation, oligomers up to nonamers were detected in the water extracted quebracho tannin.

As it was indicated earlier, the quebracho extracts exhibit similar patterns in their main content, however in the analysis of other extracts either modified or extracted by other methods, minor differences in their MALDI data were observed. For instance, the peaks observed at 772 and 904 Da indicate an acetylated dimer and

trimer respectively, and the highest intensity is detected in the extract that has undergone acid/base treatment. The acetylation of some of the –OH groups occurs as a result of this process and can be detected in small amounts [12].

In the solvent extracted and bisulphited quebracho tannins the minor series is made up of a different set of peaks which were also higher in intensity than those observed in the water extracted tannin. The peak at 964 Da for example is obtained by the elimination of an A-ring and its C4 from the 1132 Da tetramer (964-860 = 105). Also interesting are the peaks at 707 Da and 727 Da which are known to be due to degradation of the trimer at 860 Da. The difference in these peaks is 20 Da, an –OH goup is 16 Da, the additional 4 Da cannot be explained. Both these peaks are attributed to the loss of the A-ring, the 707 Da peak is only observed for the bisulphited and modified quebracho tannin whereas the 727 Da peak is observed for all the other extracts. An explanation of this phenomenon can be derived from the fact that although the fisetinidin structures are predominant some catechin units can be present and may be the cause of the additional hydroxyl group on the A-ring [17]. The peak at 685 Da that was observed to be present in higher proportion for the bisulphited extract was only present for the bisulphited and maleic anhydride modified quebracho tannins. This is in strong agreement with the findings made by Pasch et al. that modification of the tannin extracts make them more susceptible to depolymerisation reactions [12]. A clear confirmation is observed when these extracts are compared, with the bisulphited extract indicating the lowest degree of polymerisation indicating that C1 carbons in the tannin structure makes them more susceptible to hydrolysis (Figure 3.7) [12]. A closer look at the spectra reveals other peak series that are difficult to explain. Most of these peaks can be attributed to fragmentation and hydrolysis of the tannin oligomers; the resultant structures however are still unclear. This point will be further discussed in a later section.

Table 3.2: Calculated and observed masses of the solvent extracted cacao tannin (sample 25) by MALDI-TOF. The predominant repeat unit is 288 Da, corresponding to a catechin monomer.

Oligomer	Calculated $M+Na^+$	Observed $M+Na^+$		Unit type		
	Linear positive			**A**	**B**	**C**
Dimer	596.6	567		2	-	-
	601.6	602*		1	-	1
			or	-	2	-
	617.6	618		-	1	1
Trimer	889.9	891*		1	1	1
			or	-	3	-
	905.9	906		-	2	1
			or	1	-	2
Tetramer	1178.2	1178*		-	4	-
			or	2	-	2
	1194.2	1194		1	1	2
			or	-	3	1
Pentamer	1466.5	1466*		-	5	-
			or	1	3	1
	1482.5	1483		-	4	1
			or	2	-	3
Hexamer	1754.8	1754		-	6	-
			or	2	2	2
Heptamer	2059.1	2057		-	6	1
			or	1	4	2
Octamer	2331.4	2329		2	4	2
			or	1	6	1

** represents the dominant oligomer peak*

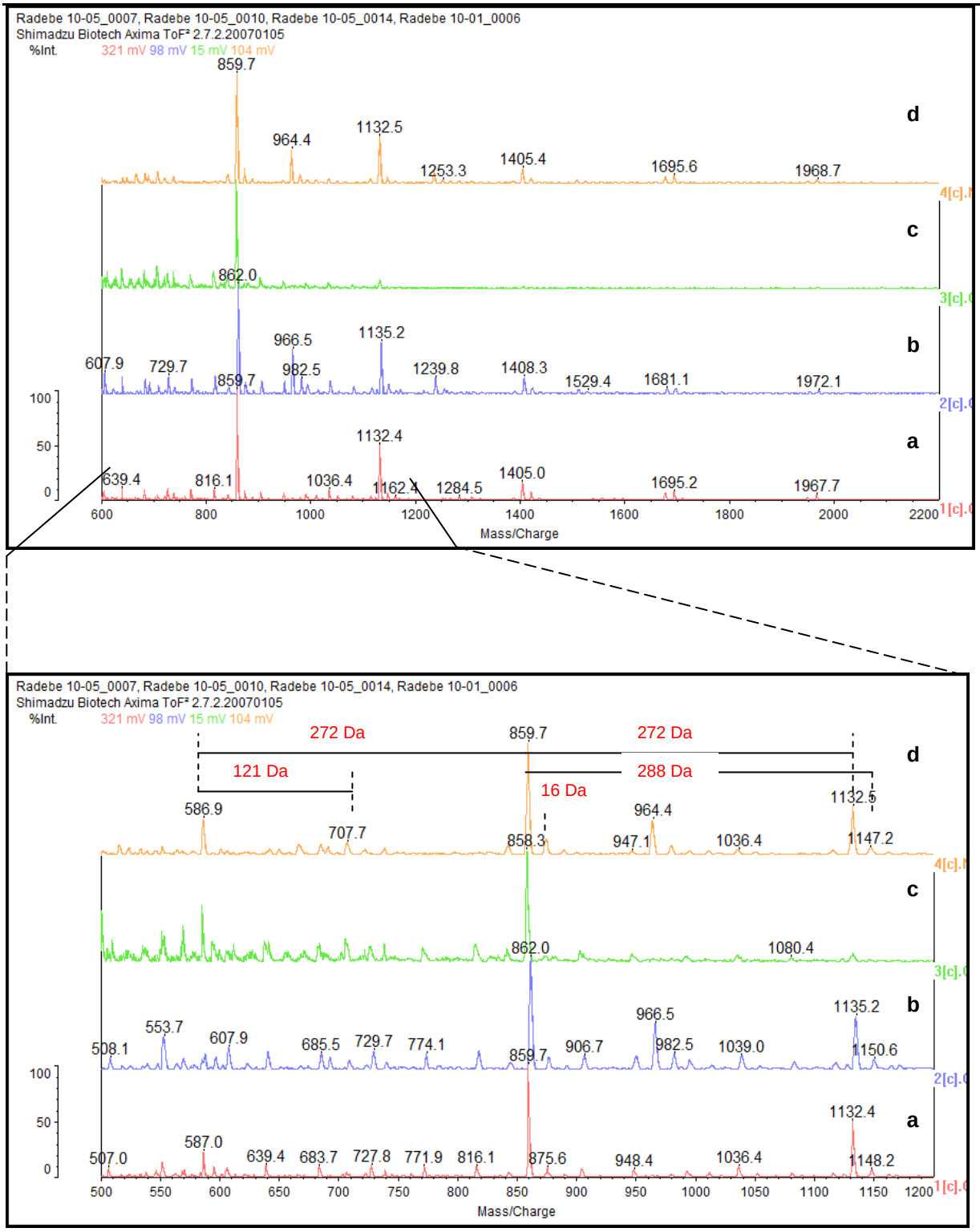


Figure 3.9: MALDI-TOF spectra of quebracho tannins (a) water extracted [sample 14] (b) bisulphited water extracted [sample 15] (c) modified [sample 16] and (d) solvent extracted [sample 12].

The main constituent of quebracho tannin is the profisetinidin monomer regardless of the method of extraction. The method of extraction and modification affects mainly the degree of polymerisation and to a smaller degree the chemical composition and end-group functionality.

Table 3.3: Calculated and observed masses of the water-extracted quebracho tannin (sample 14) by MALDI-TOF. The predominant repeat unit is 272 Da, corresponding to a robinetinidin monomer.

Oligomer	Calculated $M+Na^+$	Observed $M + Na^+$	Unit type	
		Linear Positive	**A**	**B**
Dimer	585.6	587.0*	1	1
	601.6	602.1	-	2
Trimer	841.9	841.8	3	-
	857.9	859.7*	2	1
	873.9	875.6	1	2
Tetramer	1114.2	1116.3	4	-
	1130.2	1132.4*	3	1
	1146.2	1148.2	2	2
Pentamer	1386.5	1388.9	5	-
	1402.5	1405.0*	4	1
	1418.5	1422.7	3	2
	1434.5	1438.3	2	3
Hexamer	1658.8	1661.1	6	-
	1674.8	1677.7*	5	1
	1690.8	1695.2*	4	2
	1706.8	1711.3	3	3
Heptamer	1947.1	1950.0	6	1
	1963.1	1967.7*	5	2
	1979.1	1983.0	4	3
Octamer	2219.4	2222.6	7	1
	2235.4	2239.9*	6	2
	2247.4	2256.1	5	3
Nonamer	2507.7	2512.2	7	2
		2529.1	6	3
		2542.6	5	4
Decamer	2780.0	2784.4	8	2

**dominant oligomer peak*

3.3.2.3. Mimosa extracts: polyflavonoid mass profiles determined by MALDI-TOF

In the mimosa extracts the difference between the major peaks is 288 Da. This results in this extract having the highest content of angular structures of the tannin extracts discussed. This is however not the only major distribution, from the spectra in **Figure 3.11** the structure contains various isomers in higher intensity than the quebracho and cacao extracts. The mimosa tannin was determined to contain less procyanidin units than

quebracho tannin using ^{13}C NMR [17]. Therefore its second major constituent is prodelphinidin and this is evident from the MALDI-TOF spectra. Of the three types of tannins discussed this is the only one that consists of angular trimers at 1211 Da up to pentamers first shown by Pasch et al. [12]. This extract was also shown to have the lowest degree of polymerisation by the same authors and this was confirmed by our own study. When a bisulphited and water-extracted mimosa tannin are compared the MALDI data show that there is no major difference. However the bisulphited extracted gave higher intensity in MALDI analysis, and therefore, the discussion will focus on this extract. The peak assignments are shown in **Table 3.4**.

The relative similarity in structure of both mimosa extracts also confirms the observations made by Pasch et al. that the 'linear' structure of the quebracho tannin is more susceptible to degradation than the relatively 'branched' structure and this is the reason for the minor structural change after sulphitation of this extract [12,17] . In the mimosa extract an additional distribution which is 152 Da from the major peaks is present, as previously mentioned, this mass represents the presence of galloylated structures in the extract. An example is 1315-1164 Da = 152 Da; the 1164 Da peak is a tetramer and on attachment of a galloyl unit gives the 1315 Da mass. The data for the galloylated structures is shown in **Table 3.5**. In order to calculate the theoretical mass equation was used: $M+Na^+$ = 23 (Na) +152 (galloyl, end group) + 272 A + 288 B + 304 C. Although this seems as the most possible structure, what one may consider is the possibility of some low energy fragmentations which are possible even under mild conditions. It is very difficult to form galloylated structures since a massive rearrangement needs to occur for these type structures to be present. There is another possibiity though, both the mimosa and quebracho extracts contain phloroglucinol A-rings and the heterocycle can open-up and cause the lower-end phloroglucinol ring to be released and this will have the same mass as a gallic acid residue (Figure 3.10).

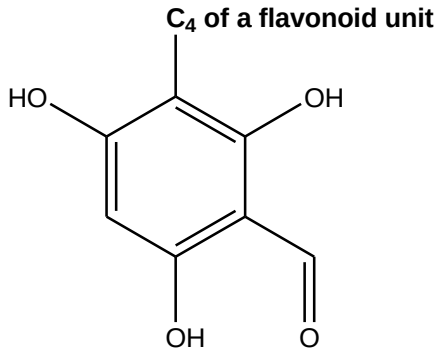


Figure 3.10: Proposed fragment obtained in the MALDI-TOF instrument with a mass of 152 Da, similar to a gallic acid unit.

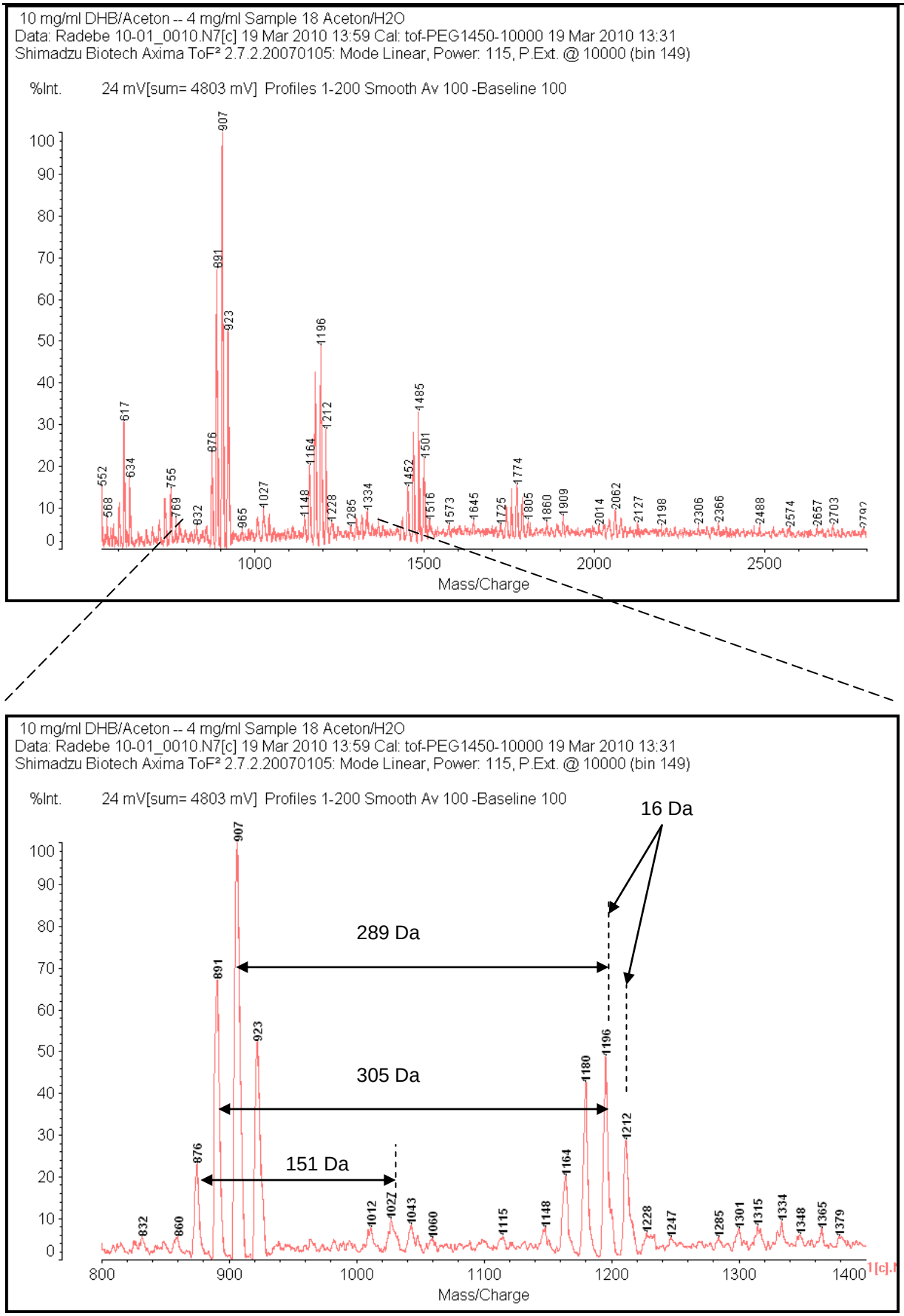


Figure 3.11: MALDI-TOF spectrum of bisulphited water extracted mimosa tannin (sample 18).

Table 3.4: Calculated and observed masses of the bisulphited water-extracted mimosa tannin (Sample 18) obtained by MALDI-TOF. The predominant repeat unit is 288 Da, corresponding to a fisetinidin monomer[a].

Oligomer	Calculated $M+Na^+$	Observed $M + Na^+$		Unit type		
		Linear Positive		**A**	**B**	**C**
Dimer	569.6	569		2	-	-
	585.6	586		1	1	-
	601.6	602		1	-	1
			or	-	2	-
	617.6	617*		-	1	1
	633.6	633		-	-	2
Trimer	857.9	860		2	1	-
	873.9	875		1	2	-
			or	2	-	1
	889.9	891		1	1	1
			or	-	3	-
	905.9	907*		-	2	1
	921.9	923		-	1	2
Tetramer	1114.2	1115		4	-	-
	1146.2	1148		2	2	-
	1162.2	1164		1	3	-
	1178.2	1180		2	-	2
			or	-	4	-
	1194.2	1196*		-	3	1
	1210.2	1212		-	2	2
	1226.2	1228		-	1	3
	1242.2	1247		-	-	4
Pentamer	1418.5	1421		4	-	1
	1434.5	1438		2	3	-
	1450.5	1452		1	4	-
			or	3	-	2
	1466.5	1469		1	3	1
			or	-	5	-
	1482.5	1485*		2	-	3
	1498.5	1501		1	1	3
			or	-	3	2
	1514.5	1516		1	-	4
			or	-	2	3

**dominant oligomer [a] data only shown up to pentamers*

Table 3.5: Calculated and observed masses of the bisulphited water-extracted mimosa tannin [Sample 18] obtained by MALDI-TOF. The data shown indicate the number of galloyl units included in an oligomer.

Oligomer	Number of galloyl units (*g*)	Number of C-type units	Calculated $M+Na^+$	Observed $M+Na^+$ (*positive linear*)
Trimer	0	0	873.9	875.5
	1	0	1025.9	1027.3
	0	1	889.9	891.4
	1	1	1057.9	1059.9
Tetramer	0	0	1162.2	1164.2
	1	0	1314.2	1314.9
	1	1	1346.2	1348.1
	1	2	1362.2	1365.2
	1	3	1378.2	1378.9
Pentamer	0	1	1418.5	1420.7
	1	2	1602.5	1605.7
Heptamer	0	2	2043.1	2045.4
	1	2	2195.1	2197.9

Mimosa extracts are less susceptible to degradation and both the method of extraction and modifications have a lesser influence on the structure of the tannin. The bisulphitation reaction replaces the –OH groups on the heterocyclic structure and thus this causes ring opening. The bisulphitation reaction is carried out intermediate to adhesive synthesis as it renders the tannin molecules more soluble in water. For details on this process refer to Chapter 2.

3.3.3. Matrix-Assisted Laser Desorption/Ionisation Mass Spectrometry Time-of-Flight Collision Induced Dissociation (MALDI-TOF CID) analyses

The MALDI-TOF spectra of condensed tannins presented in the preceding section show mass increments of 272, 288 and 304 Da, depending on the type of monomer units as discussed. As seen in the tables presented, the mass assignments may be ambiguous and direct determination of the content of a specific oligomer chain is still unknown. This has been a major problem when it comes to oligomeric tannin analysis. Fragmentation of molecules using Collision Induced Dissociation (CID) experiments can assist in providing more detailed information with regards to the specific composition of an oligomer chain [8,26]. Behrens et al. was able to use post source decay (PSD) which is a technique similar to CID in order to determine the sequence of monomer units within oligomer chains. There are various mechanisms that have been suggested to occur during CID [8,27,28]. In **Figure 3.12**, the Retro-Diels-Alder fission (RDA) mechanism is presented and in this mechanism the fragment that breaks off has a mass of 152 Da or 138 Da, this was illustrated by Rohr et al, using ESI with an atmospheric pressure chemical ionisation (APCI) interface [27]. There is another possible mechanism for fragmentation with a mass loss of 288 Da and this can be explained by the quinone-methide mechanism shown in **Figure 3.14** [27]. Considering the above, the highest intensity oligomer peak of the present samples was then selected and subjected to CID fragmentation. Three samples were analysed, the cacao tannin, solvent-extracted quebracho tannin and the bisulphited mimosa tannin. A trimer was selected in all three cases and thus the following results were obtained.

In the case of cacao tannin the m/z 889.6 ion was selected; this mass may arise from three B-type units or one A-type unit, a single B-unit as well as a C-unit (**Figure 3.14**). The fragmentation gives a relatively simple pattern with a mass difference of 288 Da as this represents a catechin unit. The fragment ion at 601.9 Da is the lowest molar mass formed and no smaller fragments are observed **(Figure 3.13)**. The 602 Da mass was also observed in the full spectrum and was determined to be composed of an A-type and C-type unit. In this case therefore it is formed by the loss of a B-type unit and thus represents a dimer, as shown in **Figure 3.14**. The peak at 738.0 Da corresponds to the loss of a single B-ring from the terminal unit by the RDA fission mechanism.

Unfortunately the peaks at 867.3, 840.8 and 815.0 Da could not be explained. Nevertheless, the main fragmentation pattern is simple enough that we can at least confirm that this oligomer is terminated by a B-

type unit. The fact that no further fragmentation occurs, hinders determination of complete oligomer sequence. In any case the structural possibility which consists of A-type oligomers could not be excluded. Due to the absence of these smaller fragments the ambiguity of assignment is still present; however the analysis did provide the information that the oligomer has a catechin terminal unit.

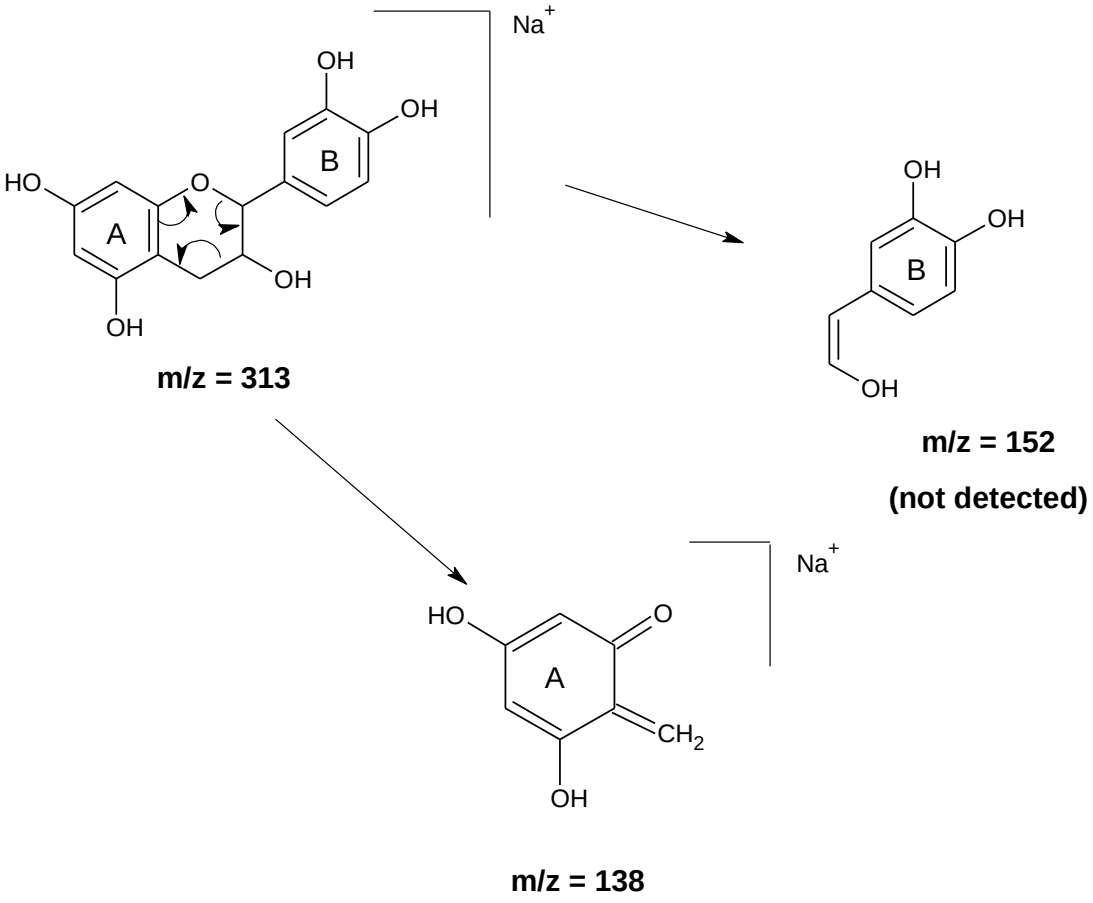


Figure 3.12: Retro-Diels-Alder (RDA) fission of a catechin monomer that occurs during MALDI-TOF CID.

The loss of 152 Da mass from the precursor ion in order to form the 738.0 Da peak can shed some light on the ring structures present in this extract. A mass of 152 Da belonging to a trihydroxylated ring, therefore this means that this oligomer has some C-type (gallocatechin) structures as well. Although the sequence could not be determined the peak at 889.6 Da in the cacao extract can be unmistakably assigned to a trimer consisting of all three monomer units; that, is one catechin monomer, a fisetinidin monomer and a gallocatechin monomer.

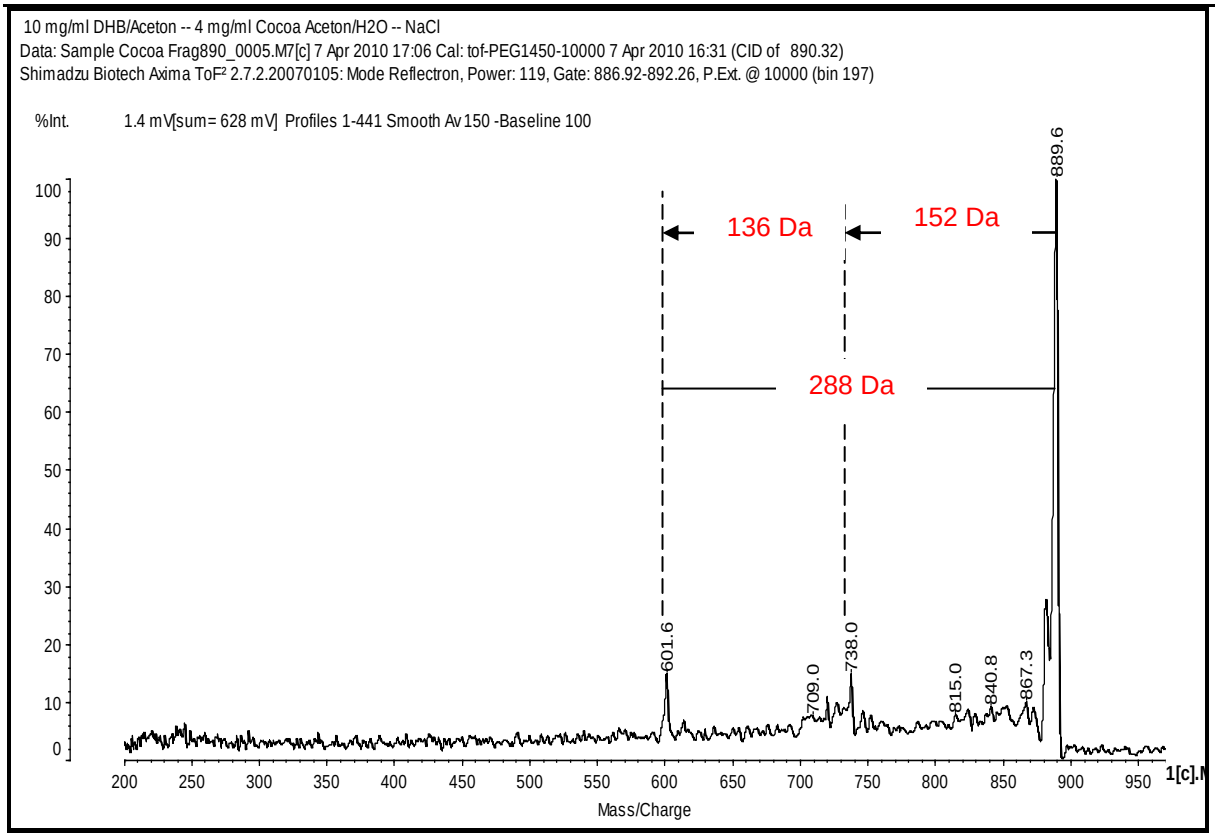


Figure 3.13: CID fragmentation spectrum of the trimer with the mass of 889 Da from the cacao tannin.

For the solvent extracted quebracho tannin an m/z 857 trimer was selected as the precursor for the CID experiment (**Figure 3.15**). The mass difference in this case is 152 Da. This is different from the pattern observed in the case of the cacao tannin; therefore the structure should be different since the mode of fragmentation differs between these two extracts. This mode of fragmentation was observed in electrospray ionisation mass spectrometry and was shown to be to a Retro-Diels Alder fission (RDA) and tends to remove the B-ring on the oligomer chain [27,28]. The RDA fission of the m/z 857 ion forms the fragment at m/z 706, which then looses a water molecule from the C3/C4 of the heterocycle to form the m/z 688 fragment ion. The loss of 152 Da mass occurs until smaller fragments are formed. In addition loss of a water molecule always seems to occur simultaneously with this mechanism. The peaks at 748 and 596 Da are also formed due to the RDA reaction; however these must be secondary products. The m/z 748 ion for example is 109 Da shifted from the precursor ion. This mass can only be explained by the loss of an A-ring from a fisetinidin monomer (A-type unit). The mechanism of this reaction is unknown to the author and as far as we know it has not been reported elsewhere. If this hypothesis is correct then it means that the structure of this oligomer can be definitively described as B-type + A-type + B-type unit, in this order. This conclusion is reached due to the fact that the loss of 152 Da can only occur at the terminus of the molecule. The 121 Da mass

difference observed between the peaks at m/z 554 and 433, could be due loss of the A-ring on the B-type unit (**Figure 3.15**).

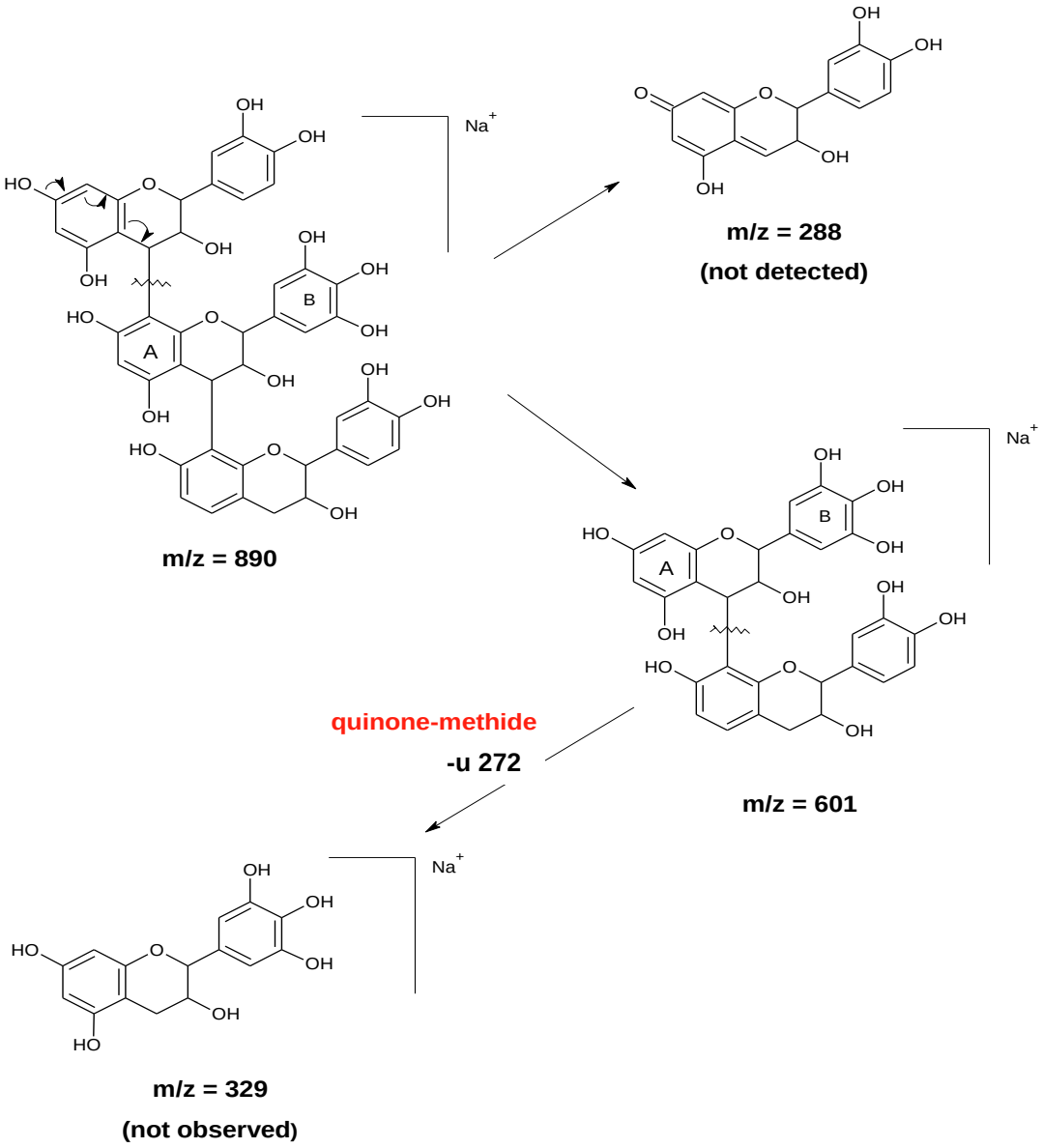


Figure 3.14: Suggested fragmentation pattern of the cacao tannin trimer making use of the quinone-methide mechanism [27].

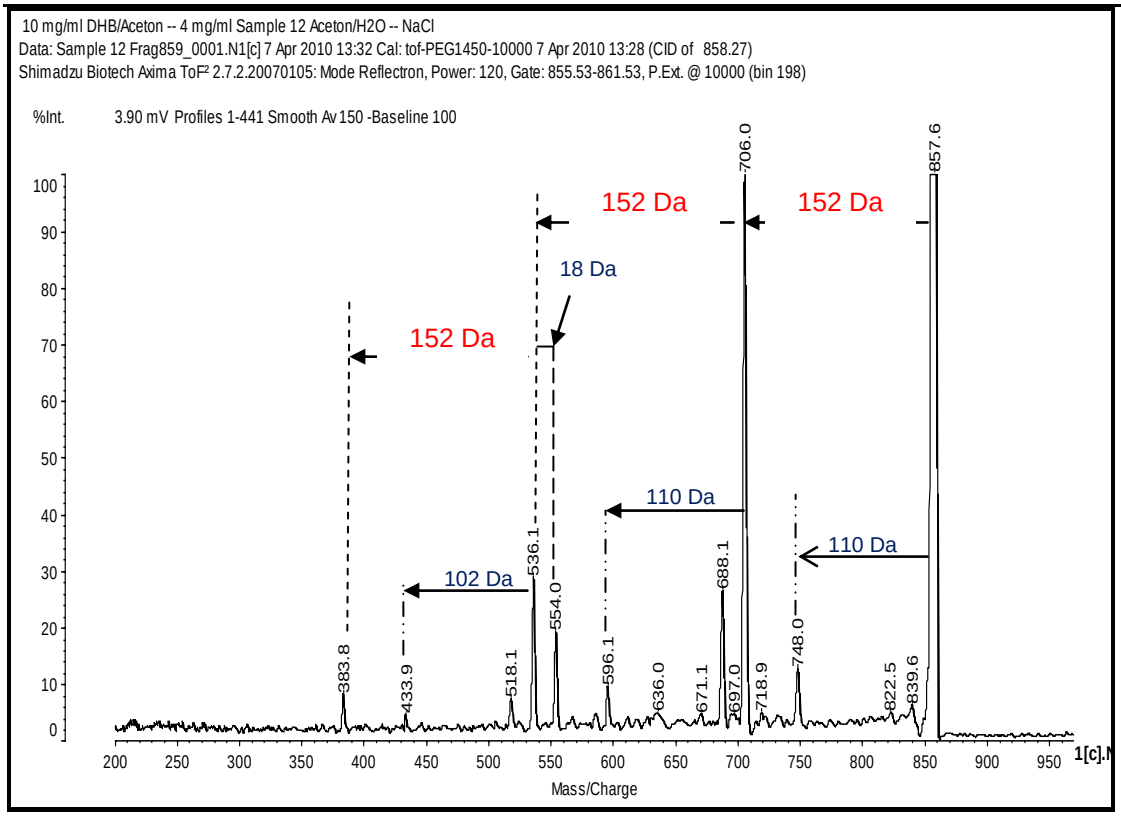


Figure 3.15: CID fragmentation spectrum of the trimer with the mass of 857 Da from the solvent extracted quebracho tannin (sample 12).

The mode of fragmentation whereby the interflavonoid bond is cleaved as in the case of the cacao tannin is not observed for the quebracho extract. Unlike in the case of ESI-MS whereby both modes are observed, MALDI-TOF seems to prefer a method depending on structure.

The same type of analysis as performed for the cacao and quebracho tannins was also carried out for the mimosa tannin. The peak at 905 Da was chosen as it was the highest intensity of the trimers. The mimosa tannin is most complex of the tannins discussed in this section. Its angular nature and numerous possibilities of structures make the mass spectral assignment complicated. The fragmentation spectrum (**Figure 3.18**) looks uncomplicated, however, the mass differences observed here are not observed for the other oligomers.

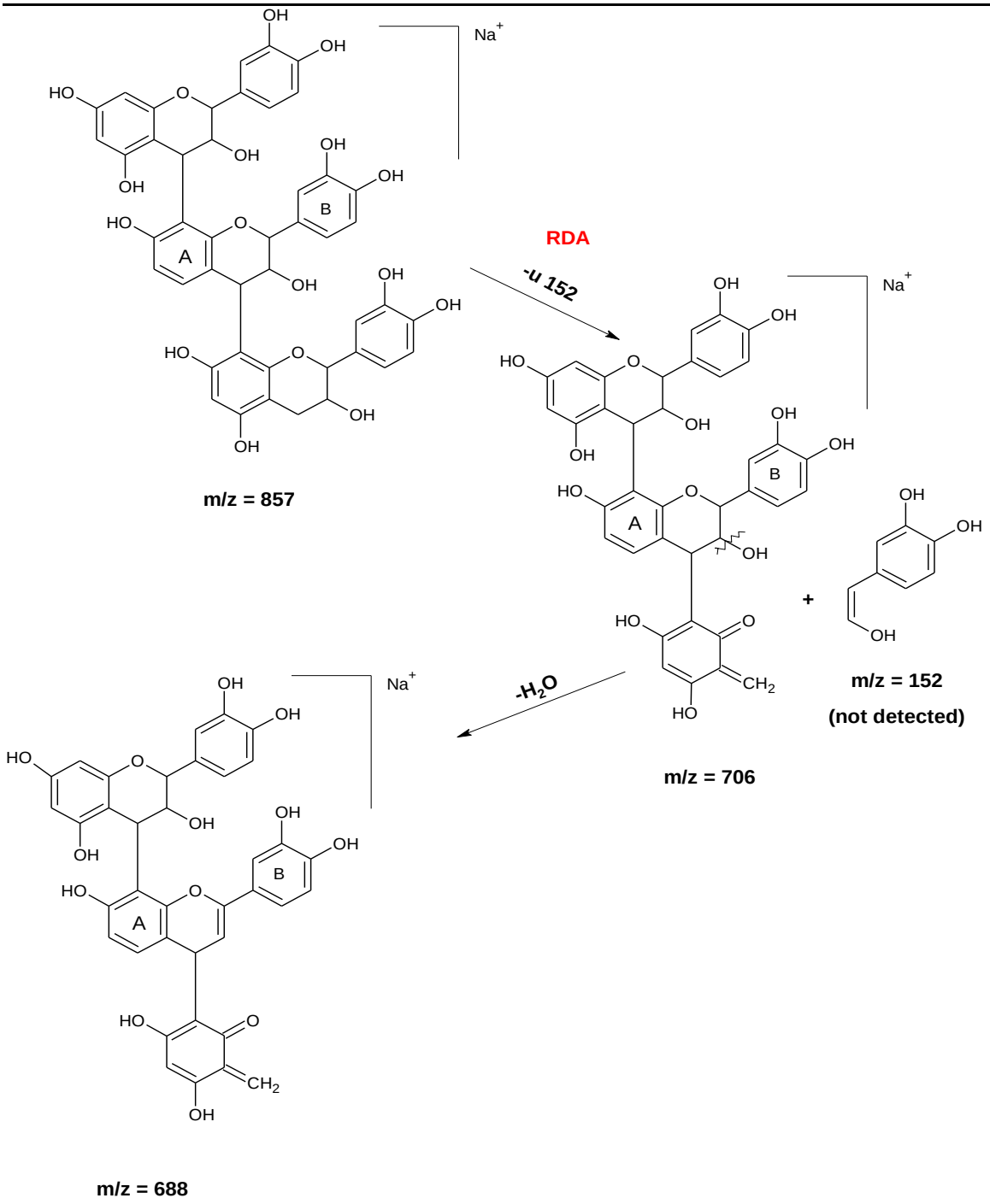


Figure 3.16: Suggested fragmentation pattern of the quebracho tannin using the RDA fission as suggested by Rohr et al. [28].

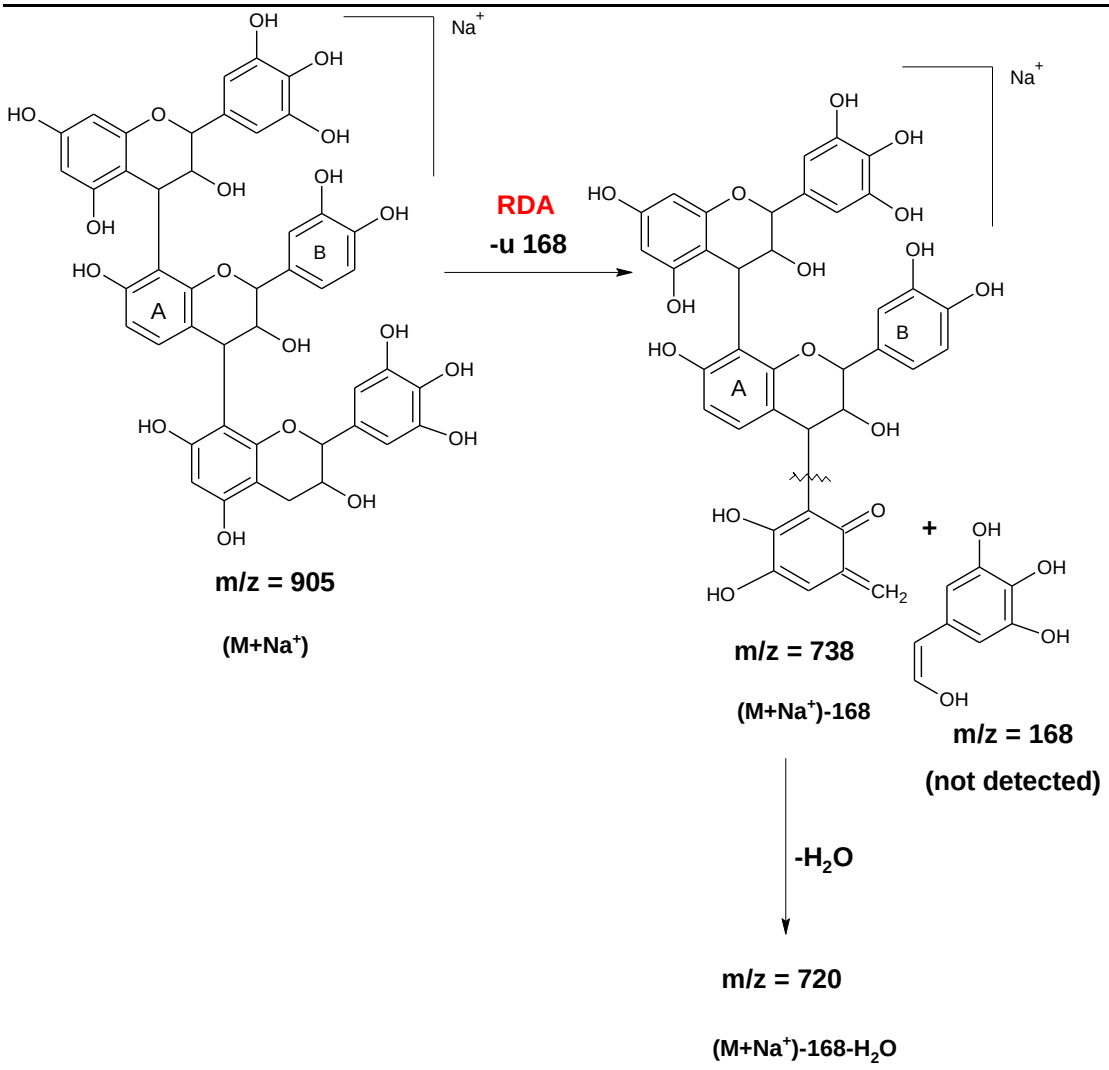


Figure 3.17: Suggested CID fragmentation pattern of the bisulphited water extracted mimosa tannin using the RDA fission as suggested by Rohr et al.[27].

In the case of the mimosa tannin the mass of 152 Da is replaced by 168 Da which clearly arises due to the presence of an additional hydroxyl group on the B-ring (**Figure 3.18**). From m/z 552 to 244 the loss is 308 Da and this clearly arises from a C-type unit. From theoretical calculation performed for the full assignment, the trimer at m/z 905 can be formed in two ways, first by two B-type units and a single C-type unit or two C-type units and an A-type unit. The loss of 168 Da from the precursor ion and also from the m/z 720 fragment ion, shows the presence of a molecule with C-type units at each terminus (**Figure 3.17**). However, the fragment that appears at m/z 685 cannot be explained by the structure that is suggested here because it is separated by 287 Da from the precursor ion. This kind of behaviour is explained by the quinone-methide mechanism and thus from this it seems the molecule also consist of a B-type terminus (**Figure 3.19**). Following this explanation then the 618 Da peak can be assigned to an oligomer that consists of one B-type

unit and one C-type unit. The mimosa tannin seems to follow the same fragmentation pattern as the cacao tannin, having both the RDA fission and quinone-methide mechanisms occurring simultaneously, however in the case of the mimosa tannin the mechanism is structure specific. Thus from this information it can be surmised that the 905 Da trimer in the MALDI-TOF spectrum is representative of two isomers which seem to be present in the mimosa extract.

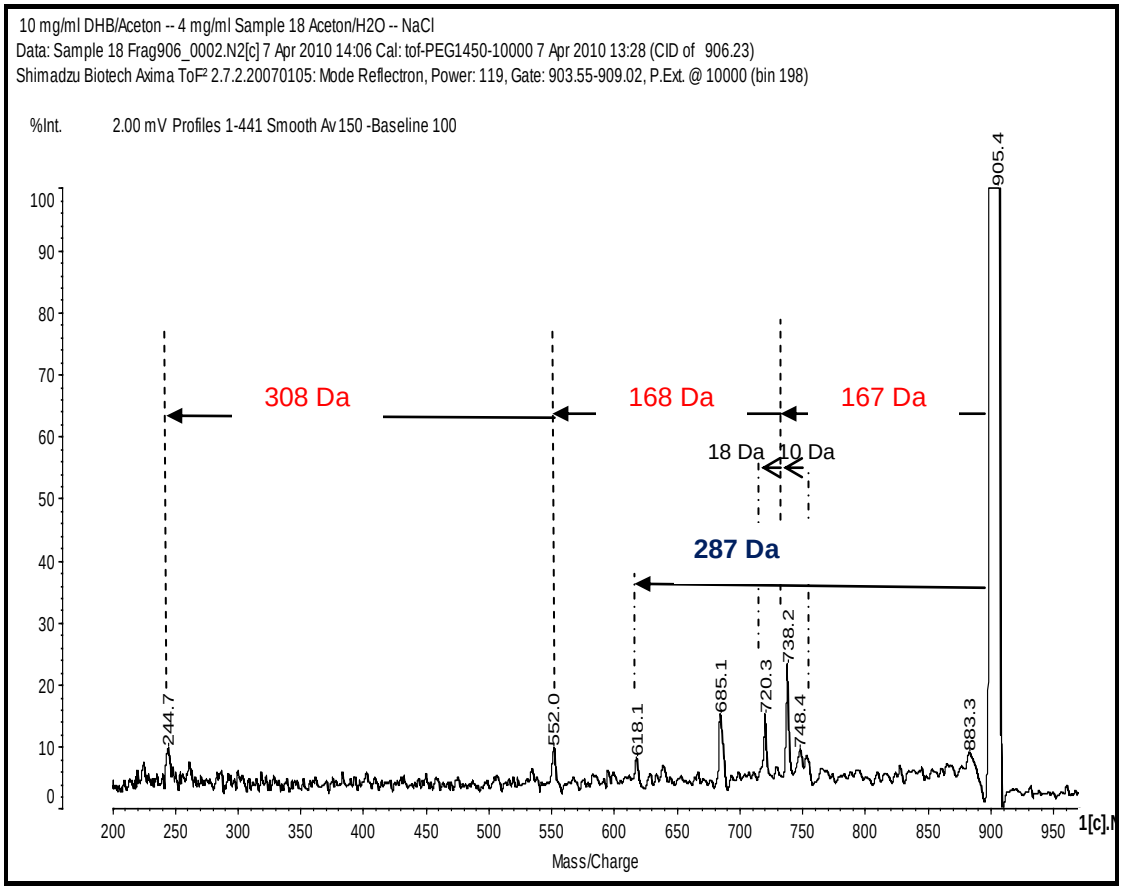


Figure 3.18: CID fragmentation spectrum of the trimer with the mass of 905 Da from the bisulphited water-extracted mimosa tannin (sample 18).

The peculiar behaviour of MALDI-TOF in fragmenting molecules with two different mechanisms was investigated further by analysing the quebracho and mimosa tannins with a softer fragmentation, namely post source decay (PSD). In this experiment the collision gas was switched off and the analysis was carried out at differing laser intensities. As can be seen in the spectra presented in **Figures 3.20 and 3.21** the fragmentation pattern is the same, however, in CID the intensity of the peaks is higher and smaller fragments are able to form. Another point is that when the laser intensity increases so does the intensity of the peaks observed in the spectra (**Figure 3.20**). It is worth noting that the laser intensities mentioned here are the

same ones normally used for tannin analysis, meaning that this form of fragmentation is inherent in the technique as well. The laser intensity on the other hand cannot be lowered in the analysis, as this gives poorly resolved spectra and in some case the analyte molecules do not receive sufficient energy to fly. This discovery means that care should be taken when assigning structures to MALDI-TOF spectra, especially in the lower molar mass range where the fragments are located.

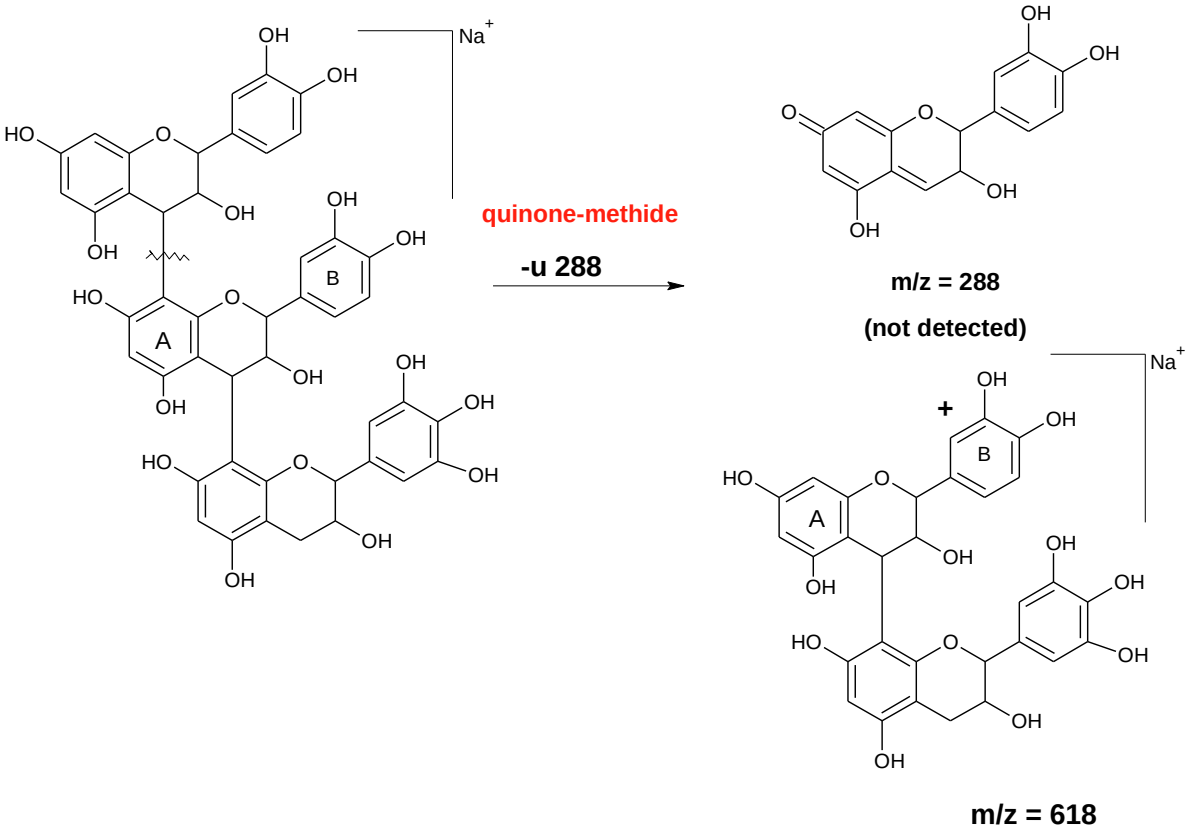


Figure 3.19: Suggested CID fragmentation pattern of the mimosa tannin trimer making use of the quinone-methide mechanism as adapted from Rohr et al. [27].

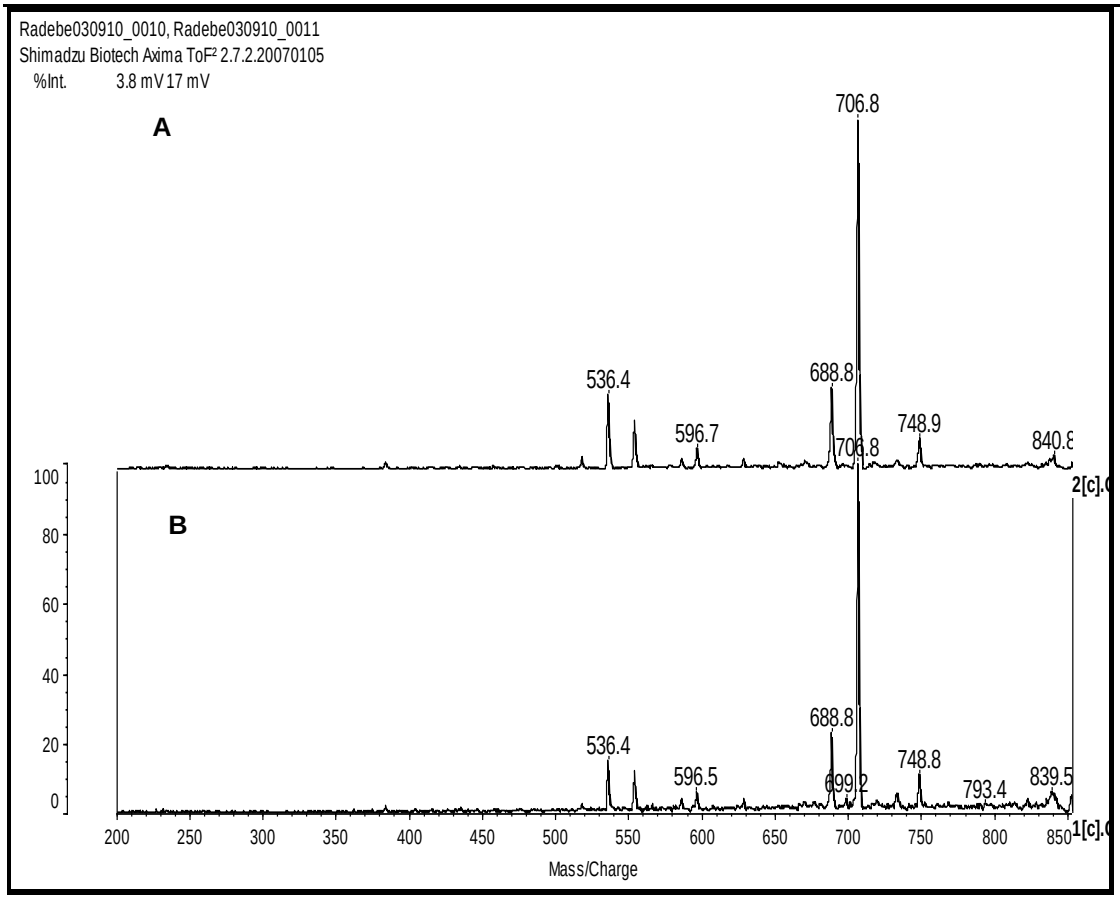


Figure 3.20: PSD fragmentation spectrum of the trimer with the mass of 857 Da from the solvent-extracted quebracho tannin (sample 12), (A) with laser power of 120 and (B) laser power of 110.

It was indicated by Pasch et al. that depolymerisation occurs during tannin extraction [12]. However, it can be argued that not all the fragments that are present are due to the extraction process since these may be artefacts of the technique. In order to determine whether the depolymerisation reactions indeed occur, a technique such as ESI-MS may be employed in the negative mode. Direct injection of the analyte molecules can give spectra that are more complex but for the purpose of deciding on this point, a detailed analysis of the spectra will not be required. Modified quebracho tannin (sample 16) was selected since it is known to be subject to these reactions, the spectrum is shown in **Figure 3.22**. The ESI-MS spectrum shows peaks that are 156 Da separated from the main oligomer distribution. These peaks are not present in the cacao tannin and thus the assumption made about the hydrolysis is correct as determined by other techniques however there is some contribution of fragments in the case of MALDI-TOF [17].

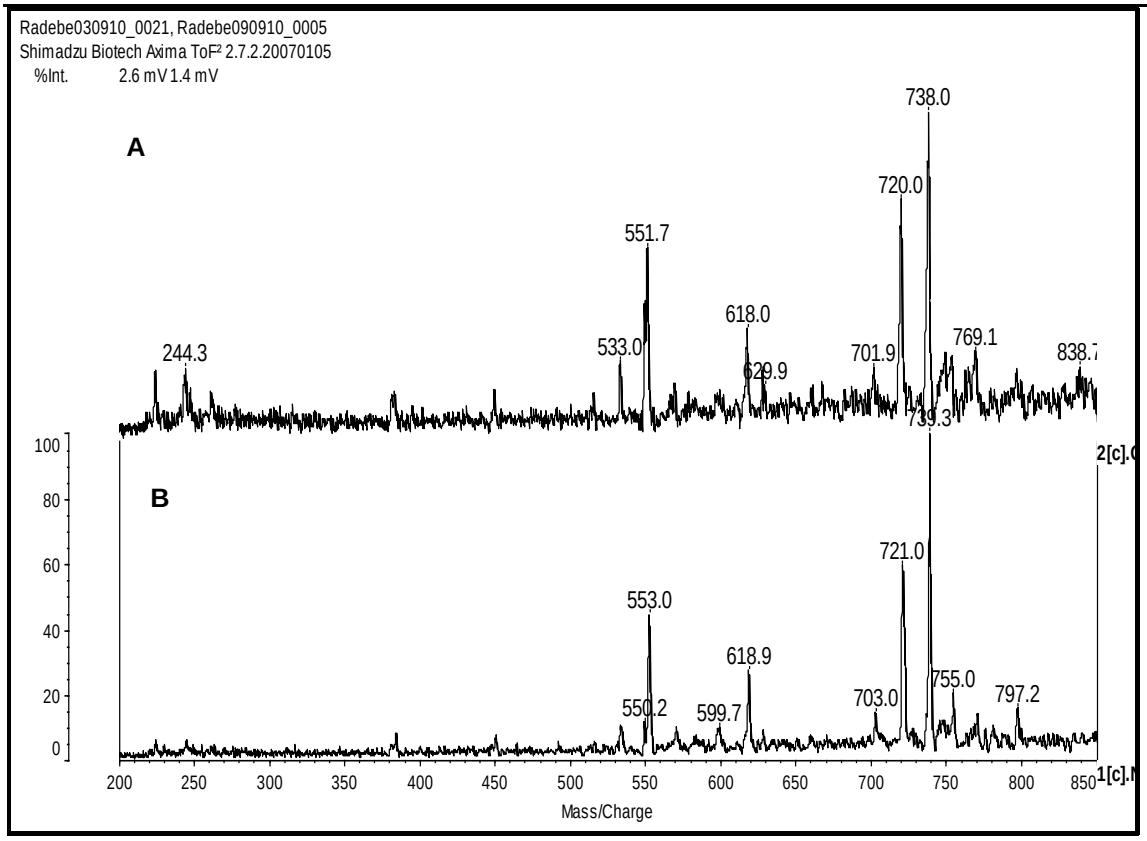


Figure 3.21: PSD fragmentation spectrum of the trimer with the mass of 905 Da from the bisulphited water-extracted mimosa tannin (sample 18), (A) with laser power of 130 and (B) laser power of 120.

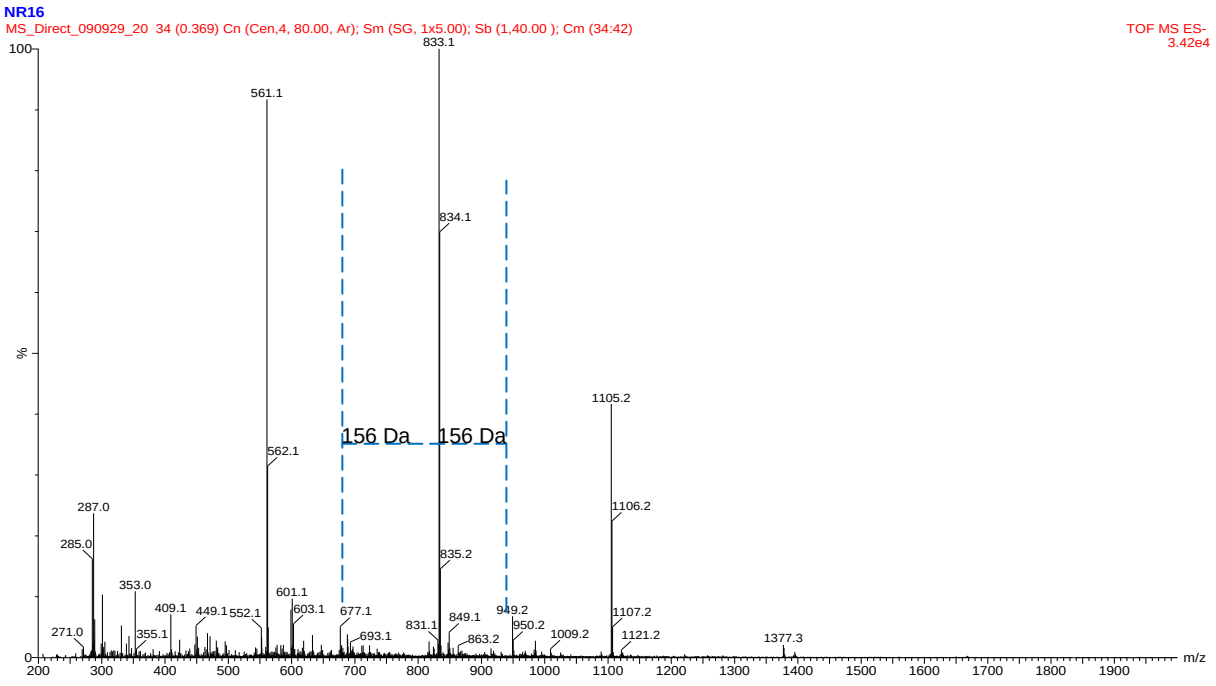


Figure 3.22: Negative mode ESI-MS direct analysis of modified quebracho tannin (sample 16).

3. 4. References

[1] A. Yanagida, T. Kanda, T. Takahashi, A. Kamimura, T. Hamazono, S. Honda, J. Chromatogr. A 890 (2000) 251.

[2] F. Pichelin, C. Kamoun, A. Pizzi, Holz als Roh- und Werkstoff. 57 (1999) 305.

[3] C. Pena, K. de la Caba, A. Retegi, C. Ocando, J. Labidi, J.M. Echeverria, I. Mondragon, J. Therm. Anal. Calorim. 96 (2009) 515.

[4] A. Pizzi, Int. J. Adhes. Adhes. 1 (1980) 13.

[5] A. Pizzi, Polymer International 39 (1996) 78.

[6] E. Haslam, Practical polyphenolics from structure to molecular recognition and physiological action, Cambridge University Press, New York, 1998.

[7] K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida, T. Yoshida, T. Matsuo, Rapid Commun. Mass Spectrom. 2 (1988) 151.

[8] A. Behrens, N. Maie, H. Knicker, I. Kogel-Knabner, Phytochemistry 62 (2003) 1159.

[9] M. Karas, U. Bachmann, U. Bahr, F. Hillenkamp, Int J Mass Spectrom Ion Processes 78 (1987) 53.

[10] F. Hillenkamp, M. Karas, R.C. Beavis, B.T. Chait, Anal Chem 63 (1991) 1193A.

[11] C.G. Krueger, N.C. Dopke, P.M. Treichel, J. Folts, J.D. Reed, J. Agric. Food Chem. 48 (2000) 1663.

[12] H. Pasch, A. Pizzi, K. Rode, Polymer 42 (2001) 7531.

[13] C. Perret, R. Pezet, R. Tabacchi, Phytochem. Anal. 14 (2003) 202.

[14] L.L. Zhang, Y.M. Lin, Molecules 13 (2008) 2986.

[15] A. Pizzi, H. Pasch, K. Rode, S. Giovando, J. Appl. Polym. Sci. 113 (2009) 3847.

[16] P. Navarrete, A. Pizzi, H. Pasch, K. Rode, L. Delmotte, Ind. Crops Prod. 32 (2010) 105.

[17] N. Meikleham, A. Pizzi, A. Stephanou, J. Appl. Polym. Sci. 54 (1994) 1827.

[18] J. Rigaud, M.T. Escribano-Bailon, J. Prieur, J.-M. Souquet, V. Cheynier, J Chromatogr A 654 (1993) 255.

[19] J.F. Hammerstone, S.A. Lazarus, A.E. Mitchell, R. Rucker, H.H. Schmitz, J Agric Food Chem 47 (1999) 490.

[20] M.A. Kelm, J.C. Johnson, R.J. Robbins, J.F. Hammerstone, H.H. Schmitz, J. Agric. Food. Chem. 54 (2006) 1571.

[21] A.L. Davis, Y. Cai, A.P. Davies, J.R. Lewis, Magn. Reson. Chem. 34 (1996) 887.

[22] A. Pizzi, D. Thompson, J. Appl. Polym. Sci. 55 (1995) 107.

[23] A. Pizzi, Wood Adhesives Chemistry and Technology, Marcel Dekker, New York, 1983.

[24] C.W. Oo, A. Pizzi, H. Pasch, M.J. Kassim, J Appl Polym Sci 109 (2008) 963.

[25] N. Vivas, M.-F. Nonier, N. Vivas de Gaulejac, C. Absalon, A. Bertrand, M. Mirabel, Anal. Chim. Acta 513 (2004) 247–256.

[26] V. Mass, W. Schrepp, B. von Vacano, H. Pasch, Macromol. Chem. Phys. 210 (2009) 1957–1965.

[27] G.E. Rohr, G. Riggio, B. Meier, O. Sticher, Phytochem. Anal. 11 (2000) 113.

[28] B. Zywicki, T. Reemtsma, M. Jekel, J. Chromatogr. A 970 (2002) 191.

Chapter 4

Multidimensional Separation and Fractionation of Oligomeric Proanthocyanidins

4. 1. Introduction

Mimosa and quebracho tannins are extracted from woods of the appropriate plant species and consist of condensed tannins. Analogous to other tannin extracts, these two tannins have distributions both in chemical composition and molar mass. The basic favan-3-ol structure which constitutes tannin has two phenolic rings, A and B-rings. The chemical composition distribution is introduced by the fact that different monomers may lead to formation of the oligomeric structure. The oligomers may be composed of catechin/epicatechin, gallocatechin or fisetinidin monomers (**Figure 4.1**) [1-3].

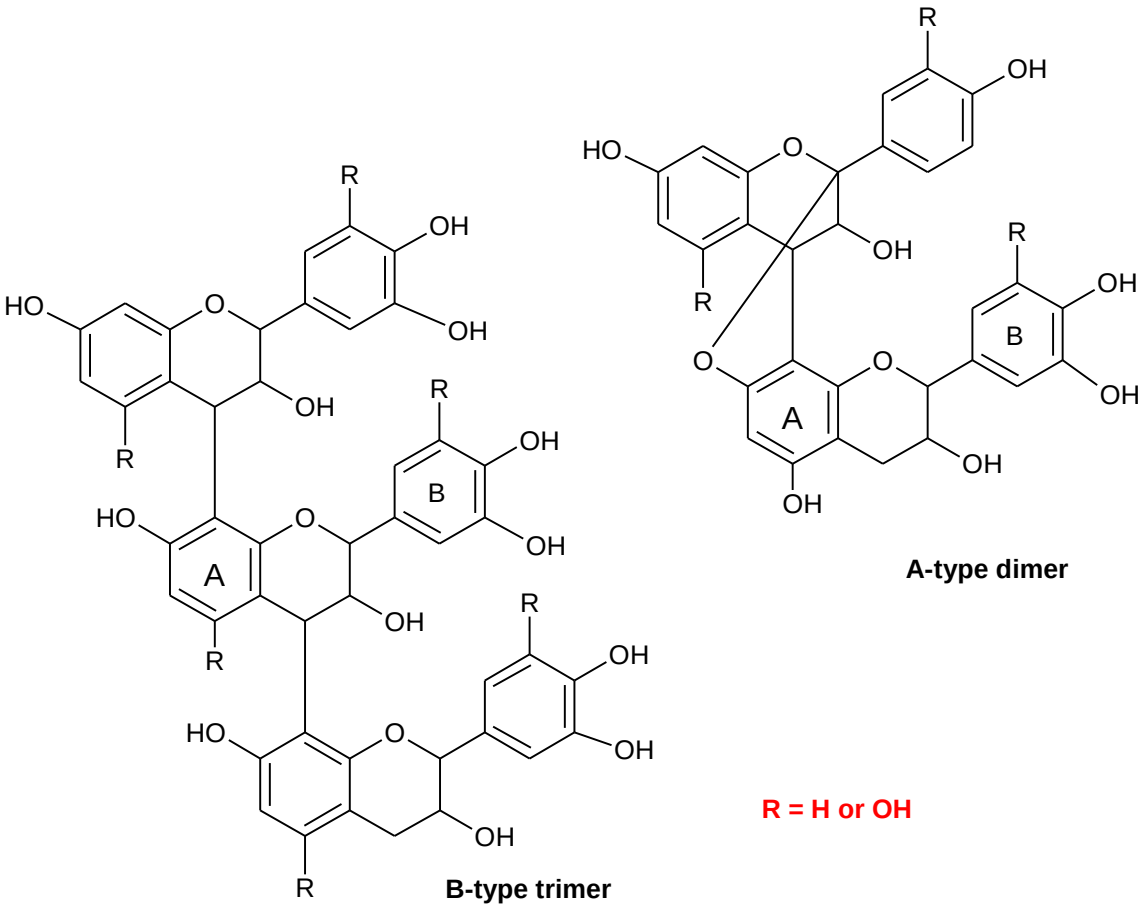


Figure 4.1: Representation of a B-type trimer and A-type dimer as may occur in a tannin extract.

The difference in these structures is basically the number of hydroxyl groups around the A and B-rings. In addition to this the monomers may be linked in different ways, either by C4-C6 or C4-C8 bonds and another

type of bond involving C4-O7 may occur [3,4]. The former bonds are more common in the mimosa and quebracho tannin and these types of bonds are called B-type, with the C4-O7 being named A-type oligomers [5]. The polymerisation of the oligomers is assumed to occur by auto condensation reactions and thus this leads to the wide molar mass ranges that are found in these extracts. Molar masses up to 20 000 units have been reported [3]. The molecules consisting of 2-6 monomer units are referred to as oligomeric and if the more units are present such molecules are referred to as polymeric [3].

Due to the wide molar mass, chemical composition and increased polarity due to the presence of the numerous hydroxyl groups along the polymer chain the analysis of proanthocyanidins is challenging. These properties of the condensed tannins also allow them to perform specialised functions which are related to their complex structure. The two extracts of interest in this case are used commercially for the synthesis of wood adhesives. The tannins are reacted with formaldehyde by condensation reactions in order to synthesise tannin-based wood adhesives. The usefulness and viability of the adhesive have shown to be affected by both the structure and molar mass. However, the structure-property relationships of tannins are difficult to determine due to the unavailability of analytical methods that are robust and selective enough to achieve complete separation and characterisation.

The purpose of this thesis was to find a chromatographic method that would enable full separation of oligomeric tannins. In order to achieve this goal, a method from literature making use of a Develosil Diol stationary phase was used since it was shown to be superior to silica columns in separation of proanthocyanidins [6]. This method showed a good separation according to degree of polymerisation and in addition some isomer separation was observed. The applicability of this method was further illustrated by Kalili and de Villiers. in its superior separation capabilities for especially higher molar mass proanthocyanidins [5]. Kelm et al. also showed a preparative scale separation making use of the same stationary phase [6]. The collection of pure fractions has been thus far a difficult task. For proanthocyanidins, preparative scale chromatography to separate the molecules by size is mostly performed on stationary phases such as Toyopearl HW-40, Sephadex LH 20 and Relite SP411 [6-10]. These methods achieve separation by means of precipitation, in that the sample is deposited and selectively dissolved using different solvents. The drawback of this method is that the pure oligomer fractions are not obtained but rather a range

of masses. Kelm et al. addressed this issue as well as the use of chlorinated solvents with silica columns by up-scaling the HILIC separation [6], discrete fractions were obtained up to heptamers.

SEC gives molar mass distribution information only, therefore in order to obtain absolute molar mass information mass sensitive detectors such as MALDI-TOF and ESI-MS are required. Ideally the separation in HILIC is by degree of polymerisation, however, some isomer separation has also been observed and therefore combination of MALDI-TOF with the separation could offer sufficient information on the type of separation obtained. The MALDI-TOF technique is also promising since it will not give multiply charged ions such as those in ESI-MS analyses, which is the more widely used technique for the mass analysis of tannins. MALDI-TOF has been combined with preparative scale liquid chromatography analysis previously, however, to the best of our knowledge no method has been published showing a HILIC-MALDI-TOF analysis of proanthocyanidins. In this thesis, a method of coupling HILIC separation to MALDI-TOF for analysis of high molar mass tannin extract from quebracho is outlined and compared to the cacao tannin.

4. 2. Experimental

4.2.1. Reagents and materials

All solvents (acetonitrile, methanol, dimethyl acetamide, dimethyl formamide, acetone and acetic acid) were of HPLC grade and purchased from Sigma-Aldrich (Steinheim, Germany). The deionised water was obtained using a Milli-Q water purification system (Millipore, Milford, MA, USA). All the crude samples were filtered through 0.45 µm nylon filter (Millipore).

4.2.2. Instrumentation

4.2.2.1 HPLC-UV analyses

Analytical scale analysis were done on an Acquity UPLC system that has a binary pump, autosampler, column oven, photodiode array (PDA) detector (500 nL flow cell, 10 mm path length). Waters Empower software (Waters, Millford, MA, USA) was used for data acquisition. In order to reduce the system dead volume, the mixing chamber of the UPLC was replaced by a low dead volume union and the tubing connected to the injector was replaced by 0.1 mm i.d. tubing.

For the increased flow rate (1 mL/min) analyses two instruments were used, the one described above and the Agilent 1200 series instrument (Agilent, CA, USA) equipped with a vacuum degasser, quaternary pump, autosampler, column oven, variable wavelength UV detector, and a semi-preparative/analytical fraction collector.

The preparative scale analysis were performed on a Waters HPLC system equipped with a binary pump, column oven, autosampler, dual wavelength detector (Waters, Millford, USA) and data acquisition was done on PSS wingpc7 software.

4.2.2.2. HPLC-ESI-MS analyses

The LC-MS analyses were carried out on the Waters Acquity UPLC system equipped with a binary solvent pump and autosampler. The LC system was connected to the Waters Ultima API quadrupole time-of-flight (Q-TOF) mass spectrometer via an electrospray ionisation (ESI) ion source. The analyses were carried out in the negative mode with a capillary voltage of 3.5 kV and cone voltage of 35 kV. The source temperature used was 100°C and the desolvation temperature was 350°C. The desolvation gas used was N_2 and the flow was 350 L/h. The gradient used is that described in the section below.

4.2.2.3 SEC analyses

Analyses were carried out on an instrument equipped with an isocratic pump (Shimadzu, Japan), autosampler, column oven, dual wavelength detector and a refractive index detector (Waters, Millford, USA). The data acquisition was performed by Millennium software. The calibration was performed using easi-vials of polymethylmethacrylate standards from PSS (Mainz, Germany).

4.2.2.4 MALDI-TOF analyses

MALDI-TOF analyses were performed on a AximaTOF^2 spectrometer (Shimadzu Biotech, Manchester, UK), equipped with a nitrogen laser (337 nm),the pulsed extraction ion source accelerated the ions to a kinetic energy of 20 keV.All analyses were carried out in the linear positive mode. Calibration was done using 1450 Da PEG standard.

4.2.2.5 ESI-MS analyses

The samples introduced for analysis on the Waters Ultima API quadrupole time-of-flight (Q-TOF) mass spectrometer by direct injection through Waters Acquity UPLC system equipped with a binary solvent manager at 300 µL/min. The capillary voltage was kept at 3.5 kV and cone voltage of 35 kV and desolvation temperature was 250°C. Analyses were carried out in the positive and negative modes. The mass range that was scanned was 100-2000 amu. Data acquisition and processing was carried out on the MassLynx v4.0 software (Waters). Calibration was done using NaF. N_2 was used as the gas and for desolvation it was kept at 250 L/h and the cone was 50 L/h. The source temperature was 80°C.

4.2.3 Chromatographic methods

4.2.3.1. Analytical scale Hydrophilic Interaction Chromatography (HILIC) analyses

The separations of the oligomeric tannins were performed on a Nomura Chemical (Aichi, Japan) Develosil Diol-100 column (250 × 1 mm i.d., 5 µm d_p,) at ambient temperature. The mobile phase consisted of (A) acetonitrile and acetic acid (99:1, % v/v) and (B) methanol, water and acetic acid (94.05: 4.95:1, % v/v/v). The separations were conducted by a series of linear gradients of B into A as follows: 4-40 % B (0-40 min), 40 % B isocratic (45-50 min), 40-4 % B (50-55 min). The analysis was done at a flow rate of 0.05 mL/min. A UV detector was used at 280 nm. For all the analysis 0.5 µL of the sample was injected in the 'partial loop with needle overfill' mode using ACN/AcOH (99:1, % v/v) as the weak needle wash solvent. Samples were prepared in a mixture of the mobile phase 25 % A and 75 % B.

For the LC-MS analysis of the oligomeric tannins a column that could reach higher flow rates was required in order to reduce analysis time and perform fewer adjustments on the instrumentation. A Nomura Chemical Develosil Diol-100 column (250 x 46 mm i.d., 5 µm d_p, Aichi, Japan) at ambient temperature. The mobile phase was the same and the gradient was adjusted in order to achieve the same type of separation at 1mL/min, for this analysis injection volume was done at 5 µL. The separations were thus effected by a series of linear gradients of B into A: 4-40 % B (0-20 min), 40-65 % B (20-25 min), 65 % B isocratic (25-35 min), 65-4 % (35-40 min).

4.2.3.2 Preparative scale Hydrophilic Interaction Chromatography (HILIC) analyses

For the preparative scale analysis, the separations were performed on a Nomura Chemical Develosil Diol-100 column (250 × 20 mm i.d., 5 µm d_p, Aichi, Japan) supplied by Separations (Randburg, SA) at room temperature. The mobile phase composition and gradient used were the same as the ones used for the analytical scale analyses. The flow rate was kept at 5mL/min and injection volume was 250 µL. Data collection was performed by UV detection at 280 nm. The fractions were collected every four minutes and the solvents were removed by rotary evaporation at a constant bath temperature of 40°C. When all the

solvent was removed 2 mL of acetone/H_2O mixture (50/50 % v/v) was added to the flask in order to redissolve the sample. The fractions were then stored in the refrigerator for varying times.

The fractions collected for ESI-MS were collected per peak using the Agilent 1200 series fraction collector. Therefore obtaining more homogenous fractions was possible. The samples were combined after 5 collections and the solvent was allowed to evaporate overnight in the fumehood. The resulting fractions were dissolved in 5 mL acetone/H_2O (50/50 %v/v) mixture. The samples were subjected to ESI-MS directly.

4.2.4. Sample preparation

Cacao beans were extracted as described by Kelm et al.[6]. Five types of commercial flavonoid tannin extracts were considered. Three quebracho (*Schinopsis balansae*) tannin extracts were measured; water extracted (**sample 14**), solvent extracted (**sample 12**), and modified (**sample 16**) quebracho tannin are outlined. The procedure for the modification of quebracho tannin for wood adhesive synthesis was outlined by Pizzi and Thompson [11]. The solvent extraction of the quebracho tannin removes the carbohydrate fraction of the extract. Two mimosa (*Acacia mearnsii*) tannin extracts from wood processed in different ways were measured, one sample is extracted by water (**sample 18**) and the other is obtained by extracting the wood with bisulphited water (**sample 19**).

4. 3. Results and discussion

4.3.1. Comparative analysis of cacao with quebracho and mimosa tannins

In order to observe the effectiveness of the HILIC method for analysis of high molar mass tannin extracts, first the cacao extract was analysed and the conditions optimised on our system. Although the primary method was introduced by Kelm et al., a shorter gradient was achieved by Kalili and de Villiers (2009) and thus this was used for further method development [5,6]. A slight adjustment of the gradient was essential in order to achieve optimum resolution for the separation. This was especially important since the high molar mass tannin extract had an unknown elution profile, a gradient that would be able to separate until 9 oligomers would suffice in the required application. The cacao tannin separation was found to be repeatable and the elution order was the same as reported in literature [6,12,13]. The same gradient was then applied to the quebracho and mimosa extracts. The UV chromatogram of the cacao is shown in **Figure 4.2**.

The high molar mass tannins have only been analysed by bulk techniques and for HPLC analyses derivitisation reactions are normally done in order to prevent any adsorption onto the stationary phase [7,14]. HPLC analyses of relatively less complex samples such as cacao have been carried out and specifically the HILIC separation has shown improved results compared to the NP separation often used for tannin analysis [6]. As mentioned, in this case the molecules are separated according to increasing degree of polymerisation. It can be thus assumed that this will be the case for the higher molar mass tannins from mimosa and quebracho, since all these extracts present similar structures. However there exist some differences in the structures and this may cause a change in the elution which is also affected by chemical composition. The assumption that the separation will be according to degree of polymerisation cannot be made directly and thus it needs to be determined definitively. For this purpose various methods will be investigated and outlined and their validities compared.

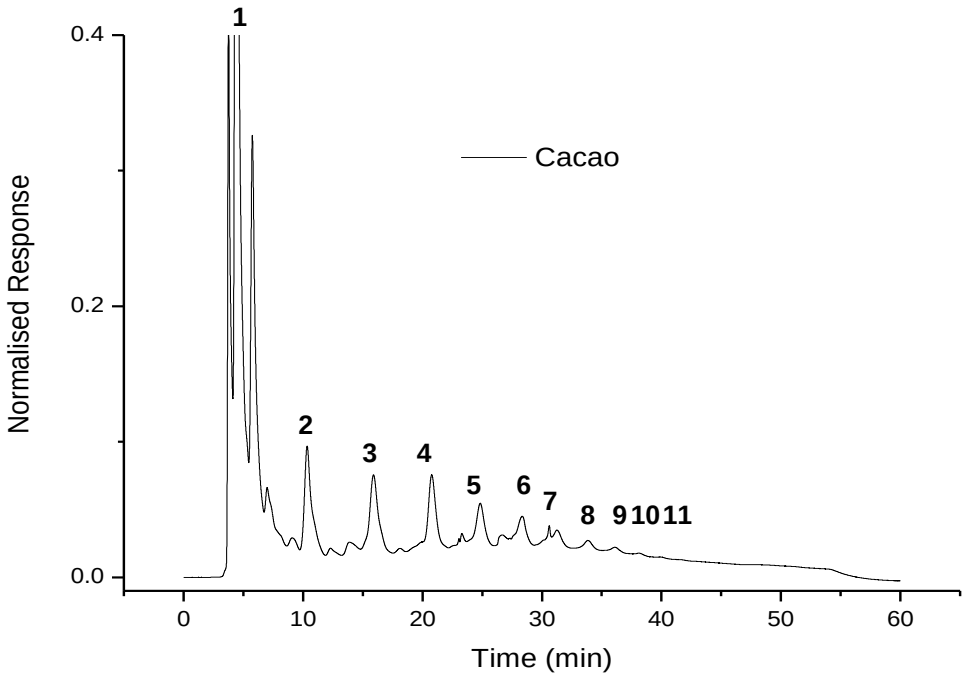


Figure 4.2: Optimised analytical scale HILIC Develosil Diol-100 column (250 × 1 mm i.d., 5 µm d_p,) of the cacao tannin . Detection: UV at 280 nm, the numbers indicate the degree of polymerisation observed for each peak.

The cacao tannin chromatogram shows 11 distinct peaks that can be assigned directly to increase in DP, whereby each number shown indicates the DP of the oligomers eluting at that point. The assignments were taken directly from literature [5,6,12,15] . The UV chromatograms of the other two tannins are shown in **Figure 4.3**, and these tannins are clearly more complex. The separation is only good at lower elution times and at higher elution time the molecules elute as a large 'hump'. This shows the samilar type of behaviour as in reverse phase separations of proanthocyanidins whereby separation can only be achieved up to tetramers. It was also indicated that when separating oligomers by size in the LAC mode the interactions may be influenced by more than the size of the molecule [9,15] . This also applies in the case of diol stationary phases which are used in HILIC separations. In the separation of tannin oligomers some isomer separation may be observed, although it is not clear to what degree this occurs.

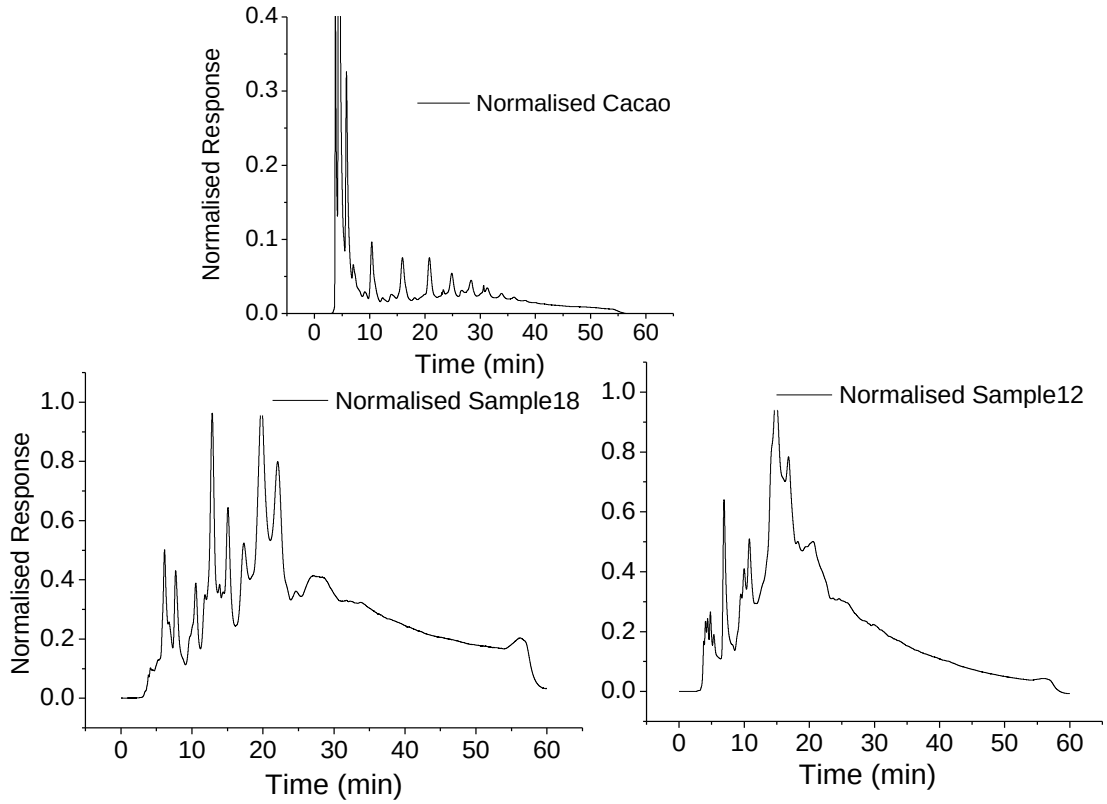


Figure 4.3: Optimised HILIC chromatograms of cacao extract, solvent extracted quebracho tannin extract (sample12) and water extracted mimosa tannin extract (sample18).Detection: UV at 280 nm.

In order to identify the peaks in the mimosa and quebracho tannins, it was assumed that the separation is by DP. By deduction, then separation could only be achieved up to heptamers i.e. 7 peaks, see **Figure 4.4**. In addition to these 7 major peaks there are shoulders on each of them and counting was done as indicated in **Figure 4.4**. When this separation is compared to the cacao tannin which has 11 peaks it is evidently incomplete, however it is clear that this tannin extract is more complex in structure and has a higher molar mass. The 'hump' observed at elution times of 35 minutes hence forth will be referred to as the 'polymeric' fraction. From this plot the direct peak assignments became more complicated since the mimosa and quebracho extracts show slightly different retention behaviour to the cacao tannin. Thus the peaks appearing in the quebracho and mimosa extracts cannot unambiguously be assigned molar mass unlike the case of the cacao extract. All the peaks from these extracts have shoulders on the major peaks and these cannot be simply dismissed since they can either be isomers of the major peaks or alternatively the oligomer with an additional monomer unit. The chromatograms in **Figure 4.4** illustrate this point; the shoulder that appears on

the dimer peak of the cacao tannin appears as a peak of high intensity in the quebracho tannin. In the case of the cacao tannin extract it was shown that the shoulders that appear on the major distribution are isomers and if this is applied to the mimosa and quebracho tannin extracts then a successful separation has been achieved up to heptamers for these higher molar mass extracts. However this inference has major flaws since all three extracts are made up of different structures.

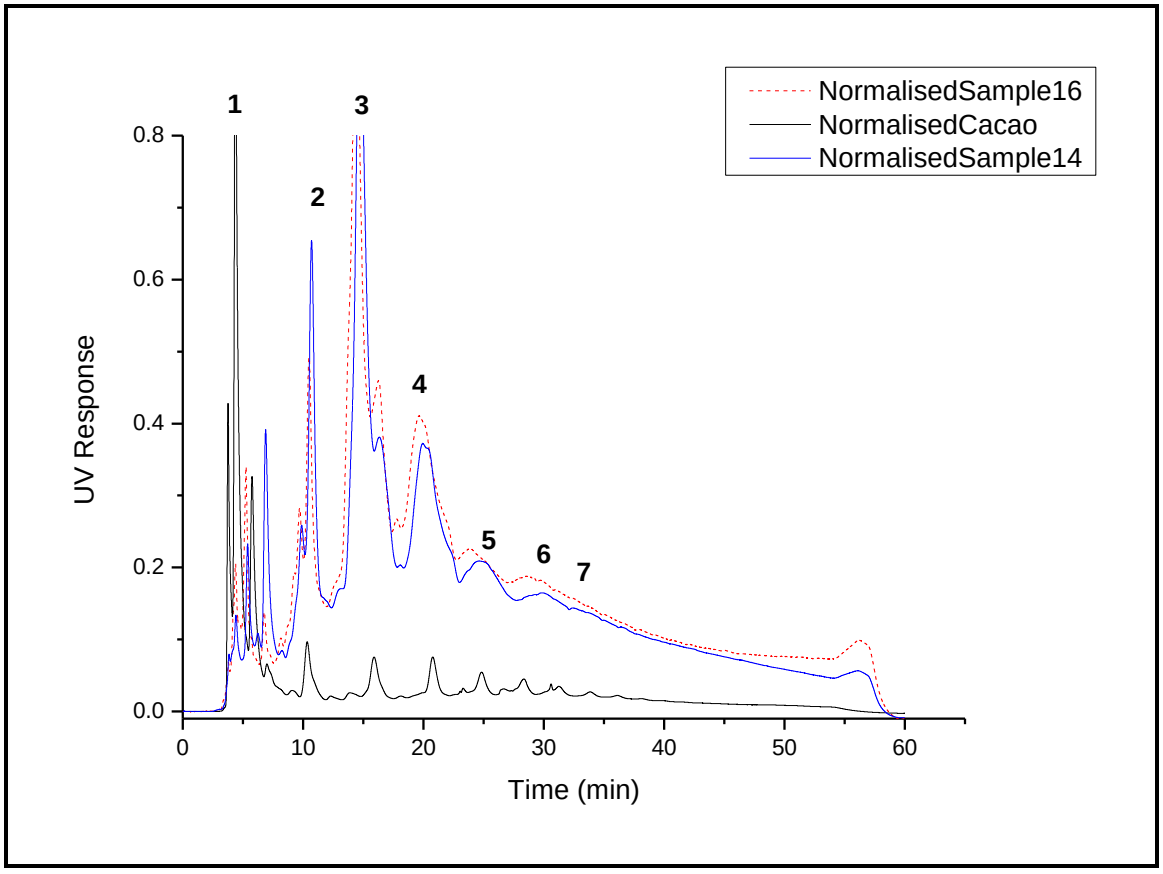


Figure 4.4: Normalised HILIC chromatograms of water-extracted (sample14) and modified (sample 16) quebracho tannins and cacao tannin. The numbers indicate the DP under each peak.

In order to establish the validity of the hypothesis stated above, a suitable detector is required that can give exact structural information. For this purpose a widely applied method is to use ESI-MS as a detector since it is able to clearly identify structures. Although LC-ESI-MS is by far the most well established method for the analysis of tannin extracts it has the limitation that it can only detect molecules that are less or equal to 5000 Da and multiple charged species are often formed during analysis. MALDI-TOF MS is the best alternative since it has also been widely used as a bulk analysis technique for the analysis of tannins and is able to

detect higher molar masses with the added advantage of the molecules being singly charged [7]. Considering these points it would then be worthwhile to analyse the molecules by LC-MALDI-TOF MS offline.

4.3.2. HILIC-MALDI-TOF-MS analysis of the cacao and quebracho tannin extracts

The chromatograms from the semi-preparative scale HILIC separation show the resolution to be the same as for the analytical scale; the results for the cacao tannin are shown in **Figure 4.5**. The analysis was performed by applying the same gradient as for the analytical scale to ensure that the separation was the same and to get good repeatability. The actual fractions were collected every 4 minutes but this meant that some of the peaks were separated between fractions. In order to address this, the appropriate fractions were combined and 5 fractions were obtained for each sample. The respective fractions were concentrated by removing the solvents by rotary evaporation. The remaining solid material was re-dissolved in a mixture of 50/50 acetone/water (% v/v), which is known to a good solvent for the tannins.

The fractions were analysed in positive mode MALDI-TOF MS. The cacao tannin was used for method development and thus the slices of the chromatogram where the fractions were taken are shown in **Figure 4.5.** The MALDI-TOF MS spectra of the earlier and later eluting fractions are shown in **Figure 4.6 and 4.7**, respectively. The resulting spectra show unusual results, when overlaid and compared to the bulk sample. Fraction 1 which is from 4-12 min (see **Figure 4.5**) is expected to contain low molar mass tannins, mainly monomers, some dimers and trimers [6,7,12]. There are numerous peaks however in the spectra and unlike the bulk sample the repeat unit of 288 Da is not observed, instead specific oligomer peaks are observed, for example the peak at 905 Da, this is a trimer with a prodelphinidin structure (**Figure 4.6**). The results from the fractions are shown in **Table 4.1**. Fraction 2 is composed of prodelphinidins and galloyted derivatives of the dimers and trimers. The first interpretable peak of this series appears at 520 Da, this peak is definitely a dimer and has three less hydroxyl groups than the first calculated dimer whose mass is 570 Da. An explanation for this peak is not evident yet, however it is related to higher molar masses by 152 Da and 304 Da, which is a galloyl and prodelphindin unit, respectively. The peak at 824 Da is related to the purely fisetinedin trimer (842 Da), they are separated by 16 Da, this means that a prodelphinidin monomer is added to the 520 Da structure in order to form a trimer. This is still perculiar since these structures are not observed

in the bulk sample. The peak observed at 474 Da represents as a glycoside of epicatechin/catechin and its presence was reported by Hammerstone et al [12].

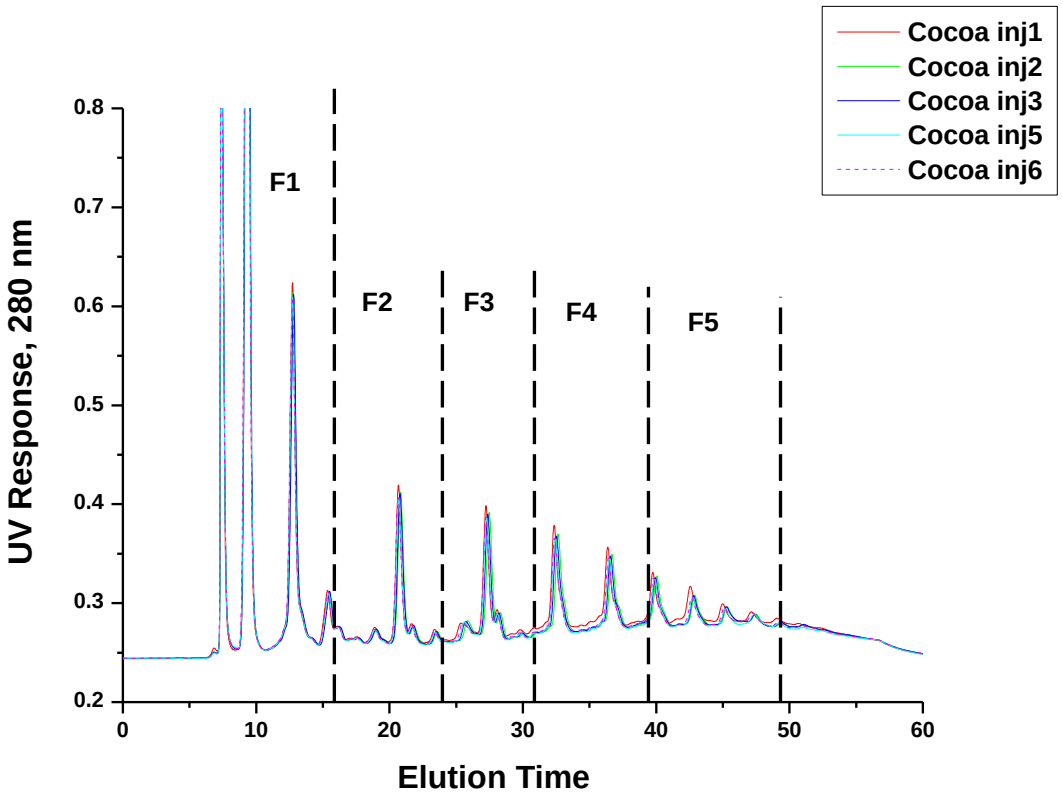


Figure 4.5: Optimised preparative scale HILIC chromatograms of the cacao tannin. The different injections are overlaid. Detection: UV at 280 nm.

Fractions 3-5 are shown in **Figure 4.7** and the spectra show definite differences in chemical composition. A closer look at all the fractions shows an increase in molar mass from fractions 1 to 5. However, the change in chemical composition along the range cannot be ignored. Fraction 4 for example, is meant to be composed of at least hexamers to octamers, however as seen in **Figure 4.7**, these are not the only oligomers present (also see **Table 4.1**). A series of peaks at lower molar mass is present and consists of a mass increment of 274 Da, which is representative of fisetinedin monomers or A-type oligomers. Fraction 5 shows hardly any peaks at higher molar masses but this could be due to the saturation of the detector from low masses such as matrix ions [7]. It is worth noting that the higher molar masses are observed only until fraction 4 as decamers (10 units). This means that the molar masses present in fraction 5 must be higher.

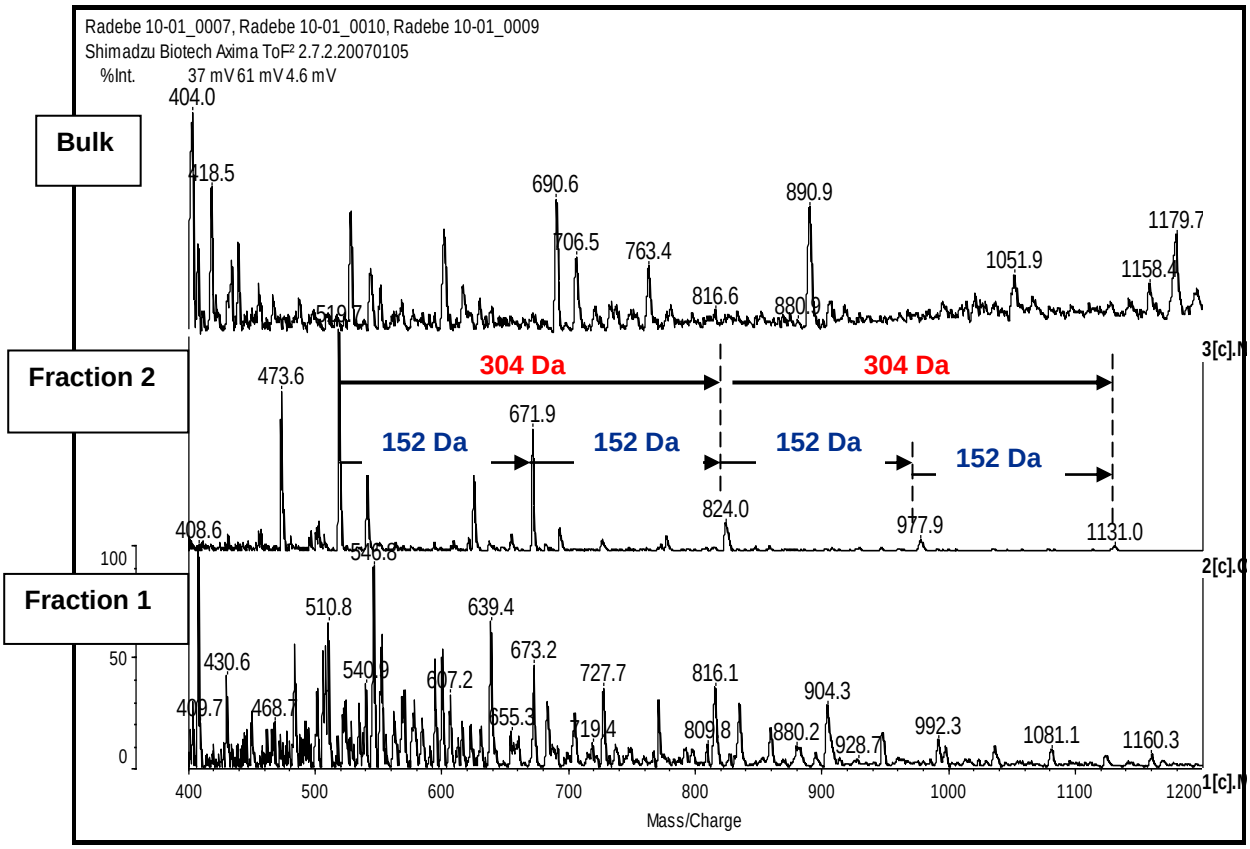


Figure 4.6: MALDI-TOF MS of fractions 1 and 2 obtained from the preparative HILIC separation of the cacao tannin. The bulk sample is also shown. The distributions present in fraction 2 are indicated.

The distributions observed in the MALDI-TOF spectra of the fractions are markedly different from the bulk sample. There are several explanations that can be given to explain this behaviour. Considering fraction 1 for example, the distributions expected are not present; this could be due to the fact that at the matrix peaks interfere with the monomer peaks fractionated. This can be corroborated by looking at fraction 2, which shows a clear distribution in molar mass, these repeat units are not observed in the bulk but could be suppressed by the higher intensity of the major oligomers in the bulk sample. Another reason for this behaviour could be due to the fragmentation of the larger molecules induced by MALDI-TOF. At the laser intensities used, it has been shown that fragmentation may occur during analysis. In fraction 2 and 3 for instance the oligomer repeat unit is 152 Da, this can either be due to galloyl units present in the extract or fragmentation of a catechin monomer induced by the laser.

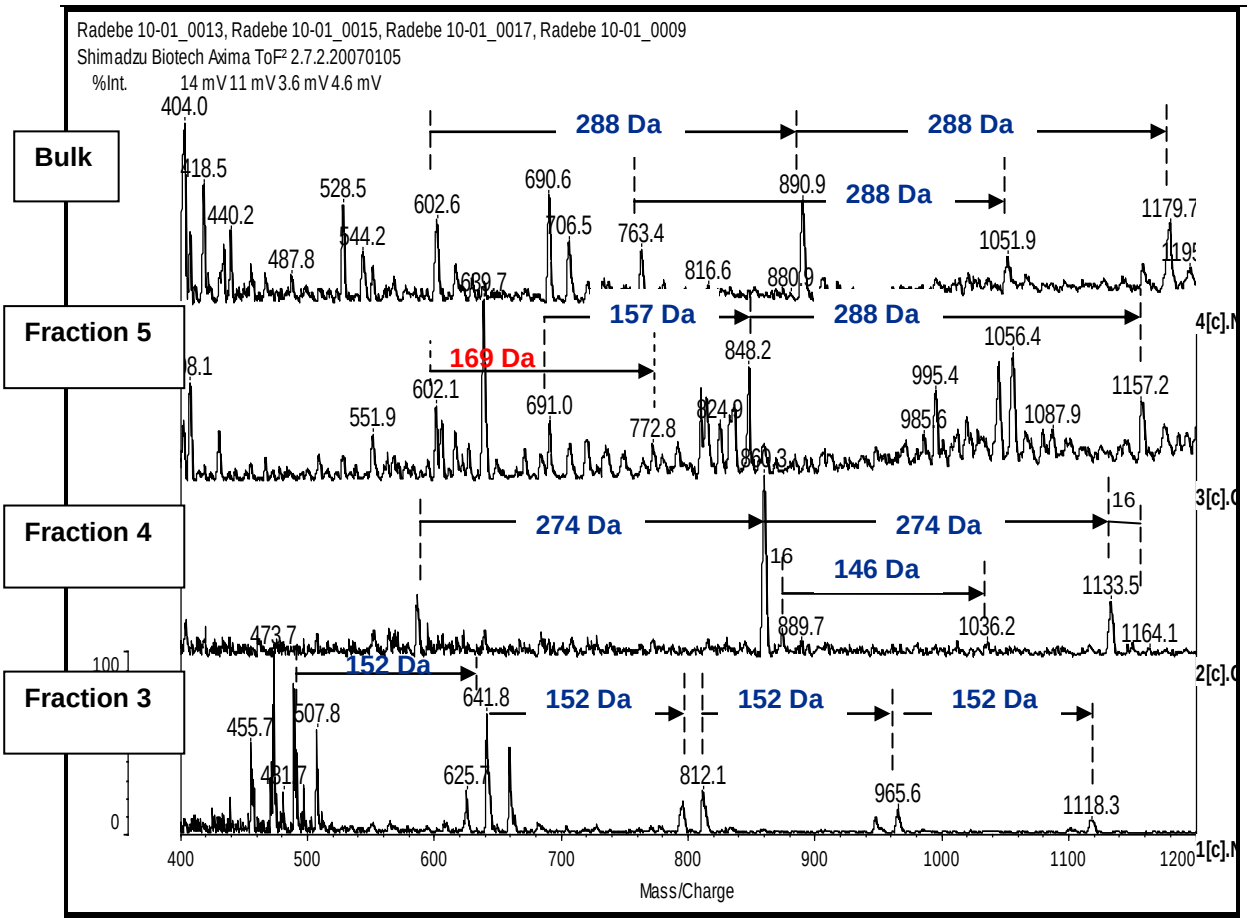


Figure 4.7: MALDI-TOF MS spectra of fractions 3-5 obtained from the preparative HILIC separation of the cacao tannin. The bulk sample is also shown.

The samples were stored for extended periods of time before actual analysis and this may have caused some hydrolysis in the oligomer structure which would then result in peaks 'disappearing' from the fractions. If however the series observed in the spectrum are actually due to the type of separation occurring in HILIC, then this would mean that the separation is also largely to according to chemical composition.

Table 4.1: Summary of the oligomer peaks observed in the MALDI-TOF MS spectra of HILIC cacao tannin fractions.

Fraction	Oligomer	Number of galloyl groups	Observed $M+Na^+$
Fraction 1 (4-12 min)	Dimer	0	585; 601; 617
	Trimer	0	860; 904
	Tetramer	0	1160
Fraction 2 (12-24 min)	Dimer	1	519
	Trimer	1	824;
	Tetramer	1	978
		0	1131
	Pentamer	1	1284
Fraction 3 (24-32 min)	Dimer	0	642
	Trimer	1	812*
		0	858*
	Tetramer	1	965
	Pentamer	1	1118
		0	1142
Fraction 4 (32-40 min)	Dimer	0	587*; 642
	Trimer	0	845; 860*; 874; 890
		1	1036
	Tetramer	0	1116; 1133*; 1164
		1	1286
	Pentamer	0	1386; 1406*; 1416; 1449
	Hexamer	0	1679; 1797;
	Heptamer	0	1952; 2010; 2092
	Octamer	0	2378
	Nonamer	0	2476; 2527; 2544; 2617;2700
	Decamer	0	2825
Fraction 5 (40-60 min)	Dimer	0	842; 858
		1	1037
	Trimer	0	1133

**dominant oligomer peaks*

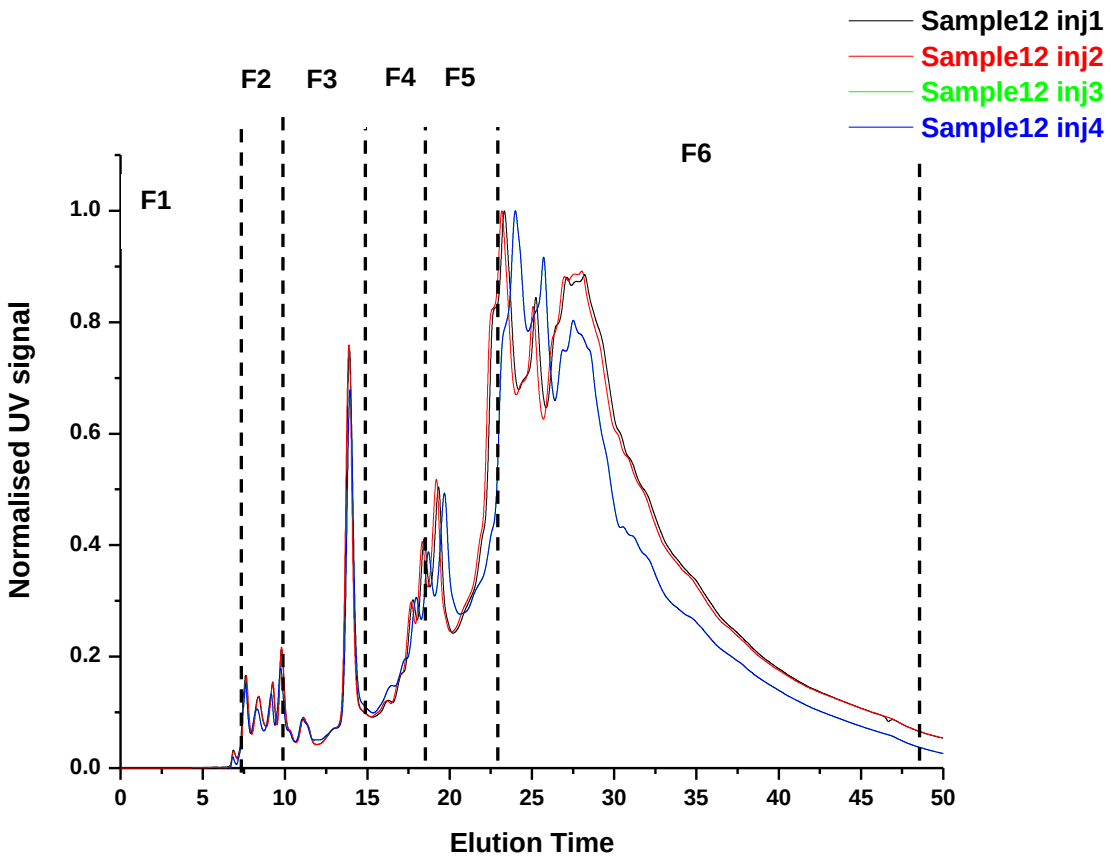


Figure 4.8: Optimised preparative scale HILIC chromatograms of the quebracho tannin. The different injections are overlaid. Detection: UV at 280 nm.

The optimised fraction collection method was applied to the solvent-extracted quebracho tannin; **Figure 4.9** shows the MALDI-TOF spectrum of the bulk sample. In this extract fraction 1 consists only of a mixture of dimers separated by a mass of 28 or 16 Da (**see Figure 4.9**) and from this point the molar mass increases with elution time as expected. The mass increment of 16 Da was observed in the bulk sample and indicates the presence of an additional –OH group on the oligomer structure. The mass increment of 28 Da observed in the fraction should be due to fragmentation processes occurring during analysis with the MALDI-TOF technique. It was shown previously that fragmentation of tannin oligomers during analysis is possible. The mass of 28 Da cannot be assigned to any of the bulk structures predicted. Considering previous experiments that were carried out, it can also be assumed that these low molar masses could arise from fragmentation due to the high laser intensity used in MALDI-TOF MS analysis. Although the matrix is able to absorb most of this energy, it has been shown that some fragmentation reactions may occur during analysis. It is not

possible to determine the source of the fragmentations just from this data, but considering that hydrolysis products were already present in the bulk it seems very likely that further fragmentation could have been induced by the high energies required for MS analysis. The repeat unit of 272 Da is observed in fraction 6, and there are some related oligomers in the less retained compounds.

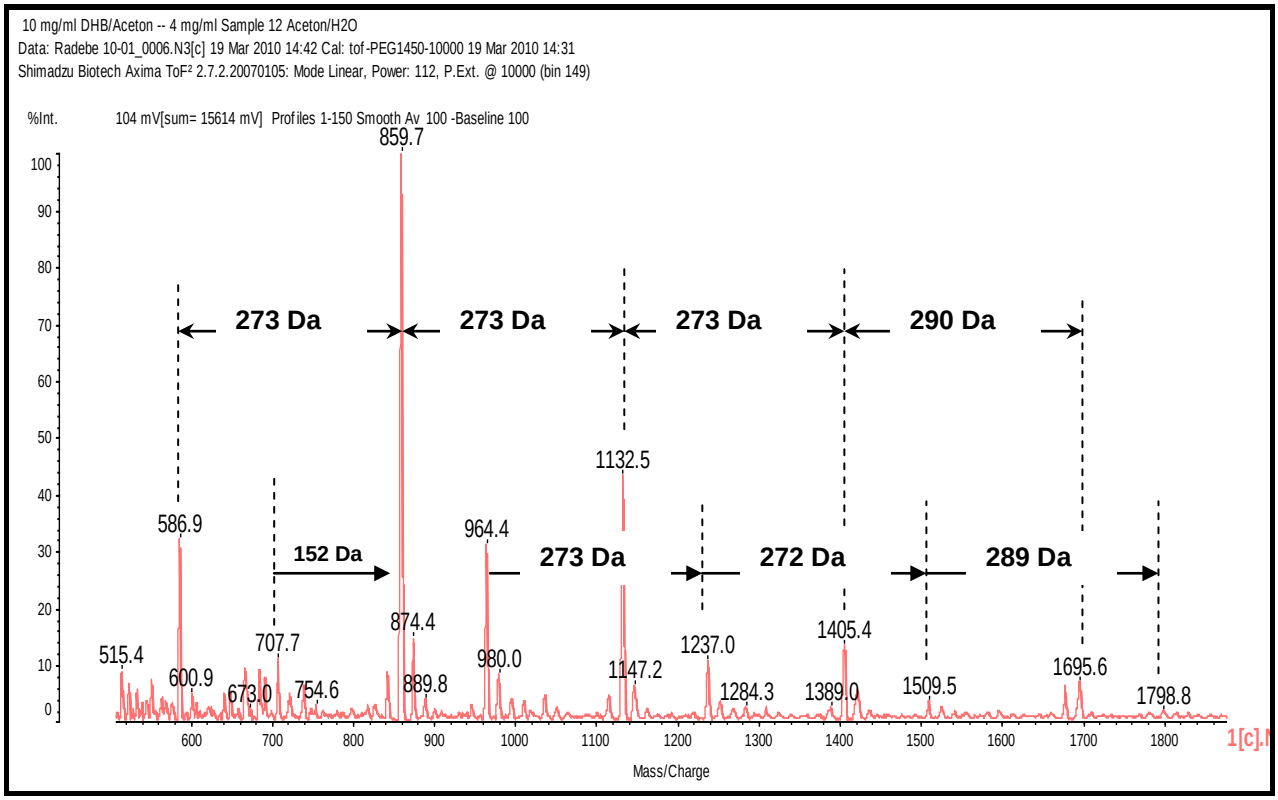


Figure 4.9: Bulk MALDI-TOF spectrum of solvent-extracted quebracho tannin (sample 12).

As in the case of the cacao tannin the polygalloyted forms of the oligomers seem to be longer retained, for instance the galloyted trimer is only observed to elute in fraction 5. In the case of the solvent-extracted quebracho tannin the galloyted oligomers are observed in fraction 5 and 6 only (**Figure 4.12 and 4.13**). In fraction 3, the peaks that are detected are the same as in fraction 1; however additional peaks with higher molar masses are also observed (**Figure 4.11**).

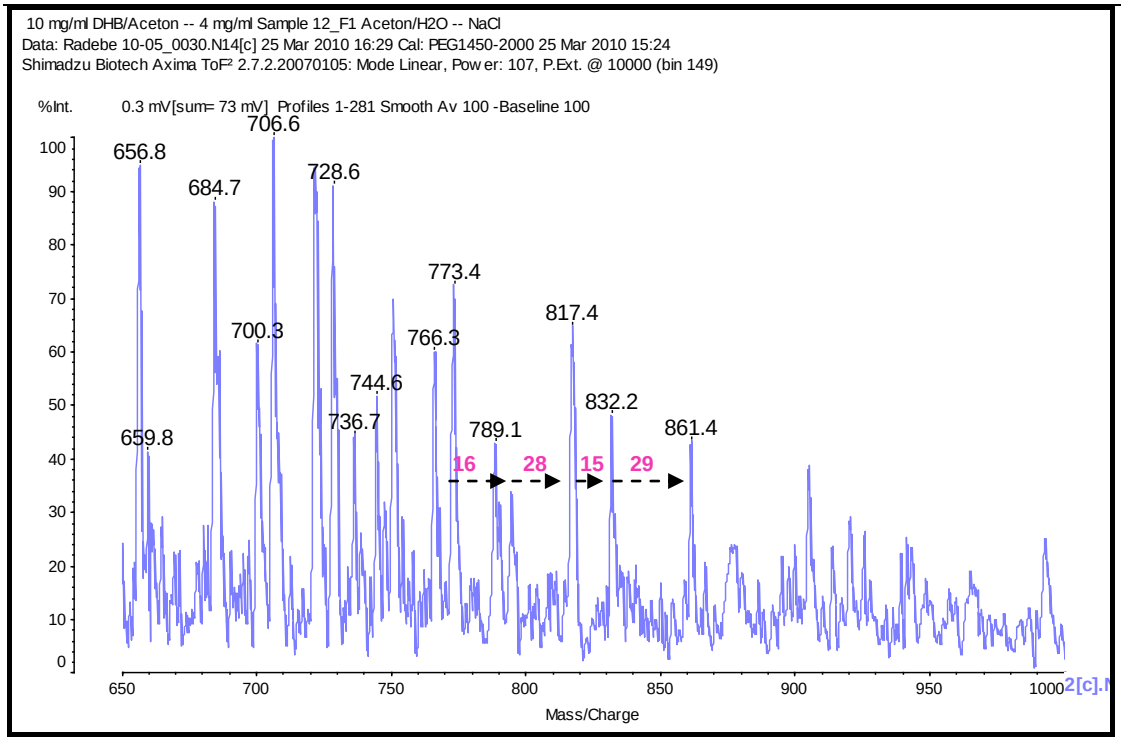


Figure 4.10: MALDI-TOF MS spectra of fraction 1 obtained from the preparative HILIC separation of solvent extracted quebracho tannin (sample 12).

Dimers and trimers are observed in fractions 4 to 6; the presence of these dimers at later eluting times suggests that there is a different separation mechanism than what was expected. For example in fraction 5 (**Figure 4.12**) only trimers are detected and unlike in the other fractions there seems to be no fragmentation of the molecules occurring. There seems to be co-elution of the low molar mass compounds with smaller molecules but this still has to be validated since there is still the question of fragment molecules that interfere with correct structure assignment. The LC-MALDI-TOF analysis shows molecules only up to tetramers; however this is not an accurate representation of the real situation since higher masses are observed in the MALDI-TOF spectrum of the bulk sample. The apparent disappearance of these higher molar masses can thus only be explained by the fact that fragmentations do occur and the masses that are observed in the MALDI-TOF spectra of the fractions are not representing completely what the sample contains. In order to check the effectiveness of this LC-MALDI-TOF method and to support some of the propositions made in this section further analysis is required that will provide the same type of information, and ESI-MS will be used for this purpose.

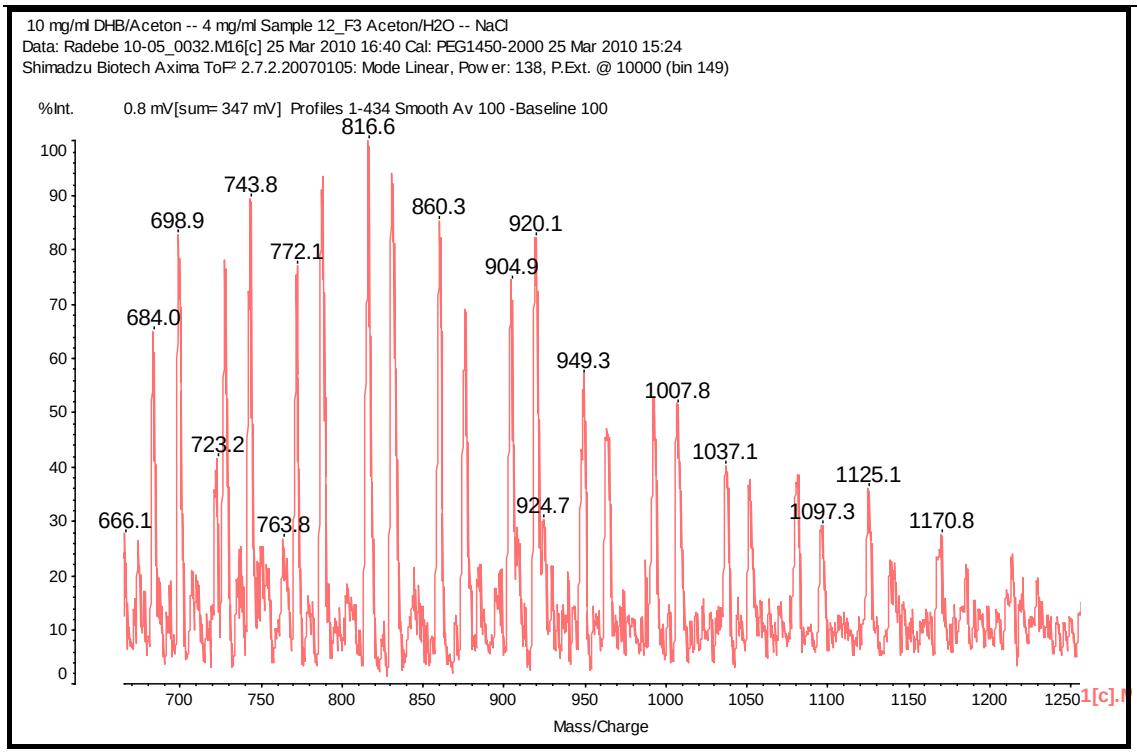


Figure 4.11: MALDI-TOF MS spectrum of fraction 3 obtained from preparative HILIC separation of solvent extracted quebracho tannin (sample 12).

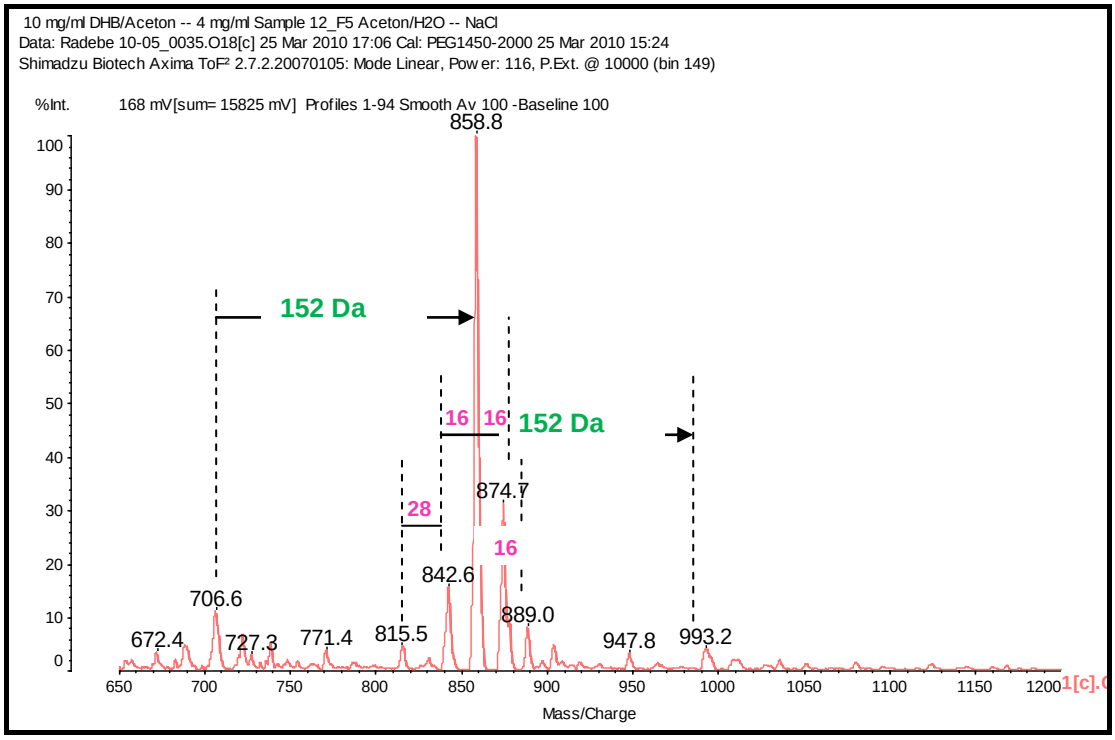


Figure 4.12: MALDI-TOF MS spectrum of fraction 5 obtained from preparative HILIC separation of solvent-extracted quebracho tannin (sample 12).

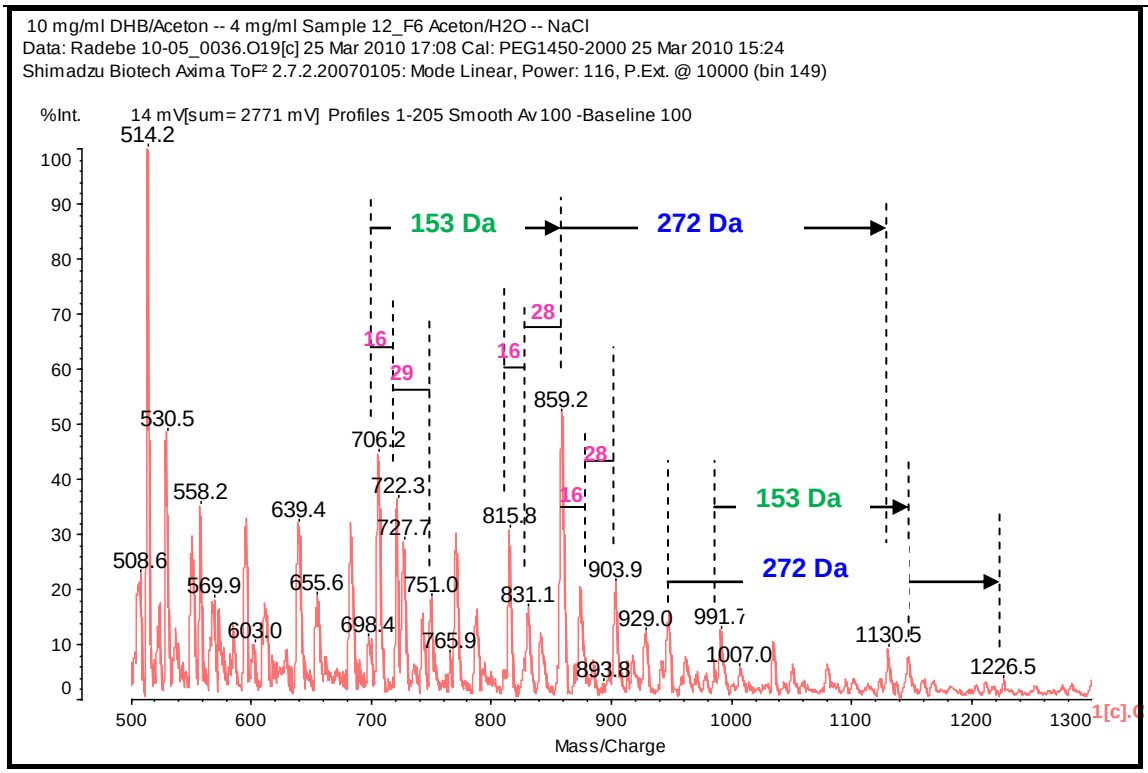


Figure 4.13: MALDI-TOF MS spectrum of fraction 6 obtained from preparative HILIC separation of solvent-extracted quebracho tannin (sample 12).

Application of mass detectors offline after HILIC fractionation revealed the elution order in this type of separation. Combination of the data obtained from LC and MALDI-TOF MS gave invaluable information in terms of the separation obtained in a HILIC separation. The 'hump' observed in the separation of the higher molar mass quebracho tannin was shown to consist of a combination of high and low molar mass constituents. The 'hump' is thus not due to the high molar mass compounds but rather due to the inability of the method to separate the isomers.

4.3.3. Analysis of condensed tannins by SEC

The SEC method described in this thesis was adapted from Kennedy et al [16]. The tannin extracts were first analysed in neat N,N-dimethyl formamide (DMF), this resulted in a great amount of tailing at lower elution volume. In consideration of this lithium chloride (LiCl) was added to the solvent and this improved the separation significantly as the tailing no longer took place. As a result of the separation at this point resembled more closely an SEC-type separation. Kennedy et al. added a small amount of water to the

mobile but this was deemed unnecessary since a good separation was achieved regardless [16]. In an attempt to avoid use of halogenated solvents dimethyl acetamide (DMAc) which has similar dissolution properties as DMF was used as an alternative.

In 'normal' SEC separations a calibration is used in order to obtain relative molar mass information of the samples of interest. The basic principle of this type of calibration works on three assumptions: (1) the separation is occurring on the basis of size (2) the molecules are present as random coils in solution (3) there is no adsorption of the samples onto the stationary phase [17]. In order to calibrate an SEC system narrow molar mass standards of the polymer of interest are required, these standards are then injected into the instrument and the elution volumes are obtained and plotted versus the Log(molar mass). The curve obtained is then applied to calculate the number average and weighted average molar masses for the polymers of interest. In reality though polymer standards of most polymers are not readily available such as in the case of tannins and therefore polystyrene (PS) standards are often used to report the relative molar masses [17]. The SEC setup used to analyse the tannins in this case used DMAc as a mobile phase on the divinyl benzene stationary phase; this combination does not allow for analysis of PS molecules to be used and thus narrow polymethyl methracrylate (PMMA) polymer standards were used for the calibration. The calibration curve obtained is shown in **Figure 4.14**. The disadvantage of using PMMA standards is that they do not give a UV signal which is the detector that is commonly used in the analysis of tannins. This limitation was overcome by using a refractive index (RI) detector which is a universal detector and can therefore be used for the PMMA standards and the tannin samples. The analysis temperature was reduced from 60°C to 40°C due to the fact that the internal temperature of the RI detector only reached up to 40°C, when higher column temperatures are used then the baseline drift is quite significant.

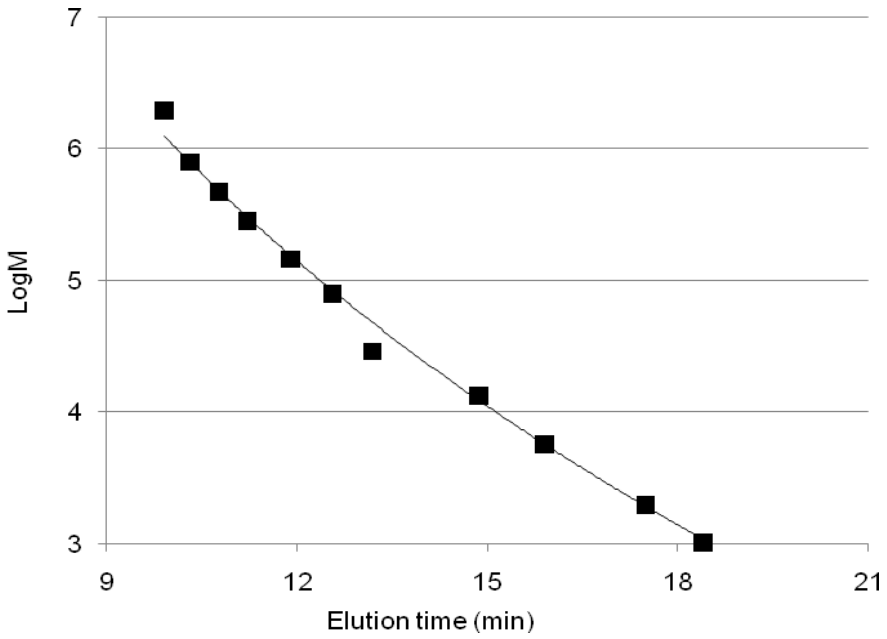


Figure 4.14: PMMA standards calibration curve used for analysis of the tannin samples.

The RI signal gave multiple peaks for all the samples which were expected since the samples have a very broad molar mass range. At lower retention times, all the chromatograms show a shoulder on the peak and this indicates the presence of higher molar masses than those reported in **Table 4.1**. The results are not in good agreement with the molar mass ranges expected for these extracts.

The M_n values of the tannin extracts vary between 5 200 and 7 500 for the major peak. These values can then be translated to give average DP varying from 20 and 28 The values are much higher than expected for these extracts and are certainly higher than those predicted by both ^{13}C NMR and MALDI-TOF [18]. This could be the result of the calibration; PMMA is not so closely related in structure to tannins and thus will behave differently in solution.

Table 4.2: Description of tannin samples and the average molar mass† determined by SEC making use of a PMMA calibration†.

Sample	Description	M_n	M_w	PDI	DP
19	Bisulphited water-extracted mimosa	5700	7400	1.2	20
18	Water-extracted mimosa	6300	10 200	1.5	22
25	Cacao	5300	10 400	2.0	18
12	Solvent-extracted quebracho	5200	8500	1.6	19
14	Water-extracted quebracho	6600	9200	1.4	24
16	Modified quebracho	7500	10 200	1.4	28

†*the molar masses are relative to narrow PMMA standards*

In SEC the molecules are separated by hydrodynamic volume and thus the conformation of the PMMA may differ to that of the tannin molecules. Unavailability of high molar mass tannin standards is a major concern and if this could be achieved it would be of interest to note the difference in the molar masses obtained by universal calibration. Another reason could be the fact that the values obtained in bulk techniques do not really represent the real situation in the sample. Due to the uncertainty, the results obtained in the SEC separation give only a very rough estimation of the molar mass information. The data of the M_n and M_w values and the samples is shown **Table 4.2** above. Quebracho and cacao tannin chromatograms are also shown in **Figure 4.16** to illustrate the robustness of the separation in SEC using these conditions. According to the SEC data, the cacao extract contains the lowest molar mass as expected. Another point to note from this chromatogram is the absence of the shoulder present in the other extracts. The building units in the cacao extract have higher molar mass compared to those that make up the quebracho extracts. This is a clear indication that separation in SEC is exclusively according to molar mass and chemical composition does not play a role in the separation. The elution profile of sample 12, which is the solvent extracted tannin, is narrower than its counterparts. Additionally this narrow profile also has the lowest molar mass as

confirmed by bulk techniques. This proves again the theory that under solvent extraction some of the tannin oligomers undergo hydrolysis and thereby reducing the molar mass. Although the molar mass is lower, a slightly bimodal distribution is observed (**Figure 4.23**).

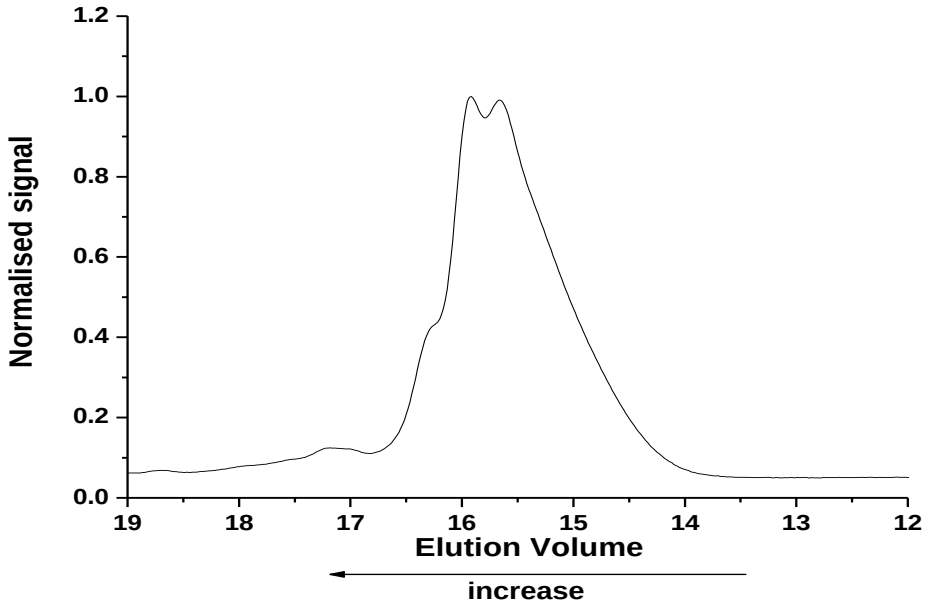


Figure 4.15: Normalised RI detector signal of an SEC separation of solvent-extracted quebracho tannin (sample 12). Stationary phase: 2 x PLgel mixed- D columns (300 × 7.5 mm i.d., 5µm) connected in series and protected by a guard column with the same material. Mobile phase: DMAc with 0.03 % (w/v) LiCl and 0.05% (w/v) BHT. Flow rate: 1 mL/min.

In addition, the cacao extract has the simplest structures, since it is made up catechin/gallocatechin building blocks whereas in the quebracho and mimosa extracts other structures are present which may have lead to the conclusion that they are higher in molar mass. The complexity added by the extraction methods may lead to a change in structure and thus make it more difficult to separate the molecules in HILIC separation. The modified quebracho extract has the highest molar mass but the degree of polymerisation seems to be the same regardless of the extraction method. This added complexity is the major reason for the apparent higher molar mass of the commercially extracted tannins.

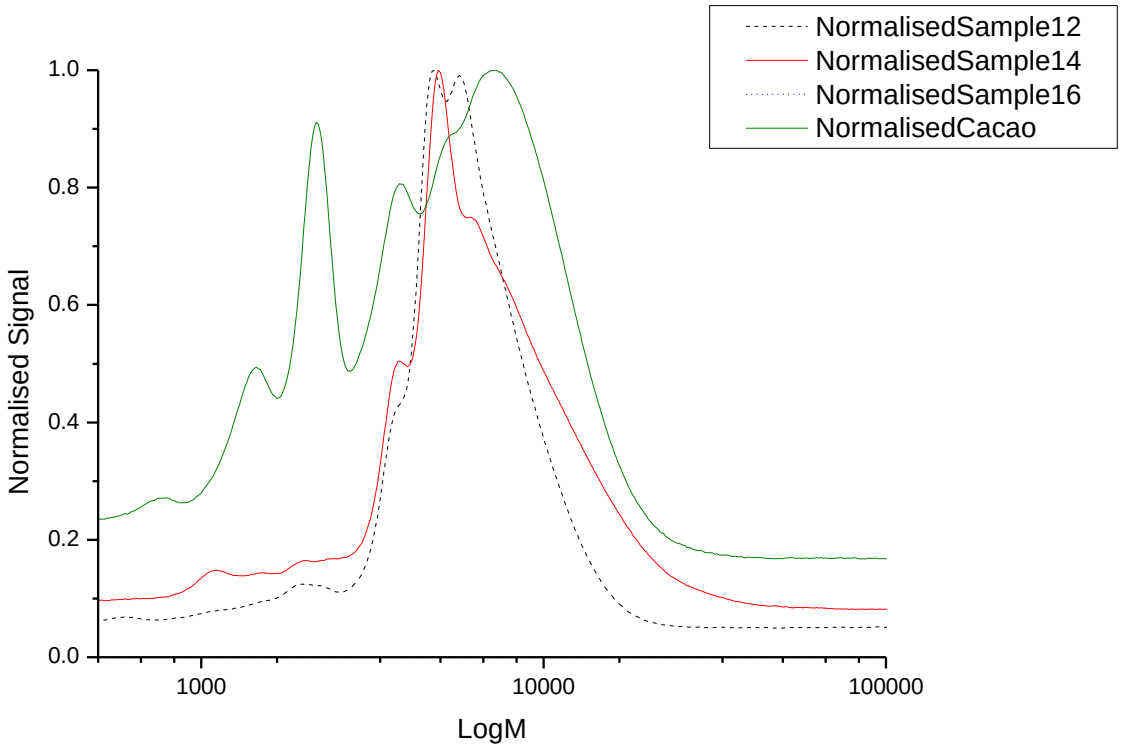


Figure 4.16: Signal response vs. LogM plots obtained by SEC of quebracho tannins, samples 12, 14 and 16 overlaid with cacao tannin, calculated using the PMMA calibration. For sample details refer to Table 3.1. Stationary phase: 2 x PLgel mixed- D columns (300 × 7.5 mm i.d., 5µm) connected in series and protected by a guard column with the same material. Mobile phase: DMAc with 0.03 % (w/v) LiCl and 0.05% (w/v) BHT. Flow rate: 1 mL/min, detection: RI.

The mimosa extract was also analysed under the same elution conditions and the molar mass distributions of the bisulphited and water extracted sample are shown in **Figure 4.17**. The elution profiles are similar with the elution of the bisulphited sample being shifted to slightly higher masses. Unlike the quebracho tannin, mimosa seems to be less susceptible to hydrolysisl, in that the molar mass was not reduced as a result of the extraction method. This could be owing to its angular structure as observed by Pasch et al. [18]. The elution profile of the mimosa tannins is not that different to that observed for the quebracho extract this is also indicative of the similarity in structure. The application of this SEC method has shown that a direct separation of oligomeric and polymeric proanthocyanidins solely by size is possible.

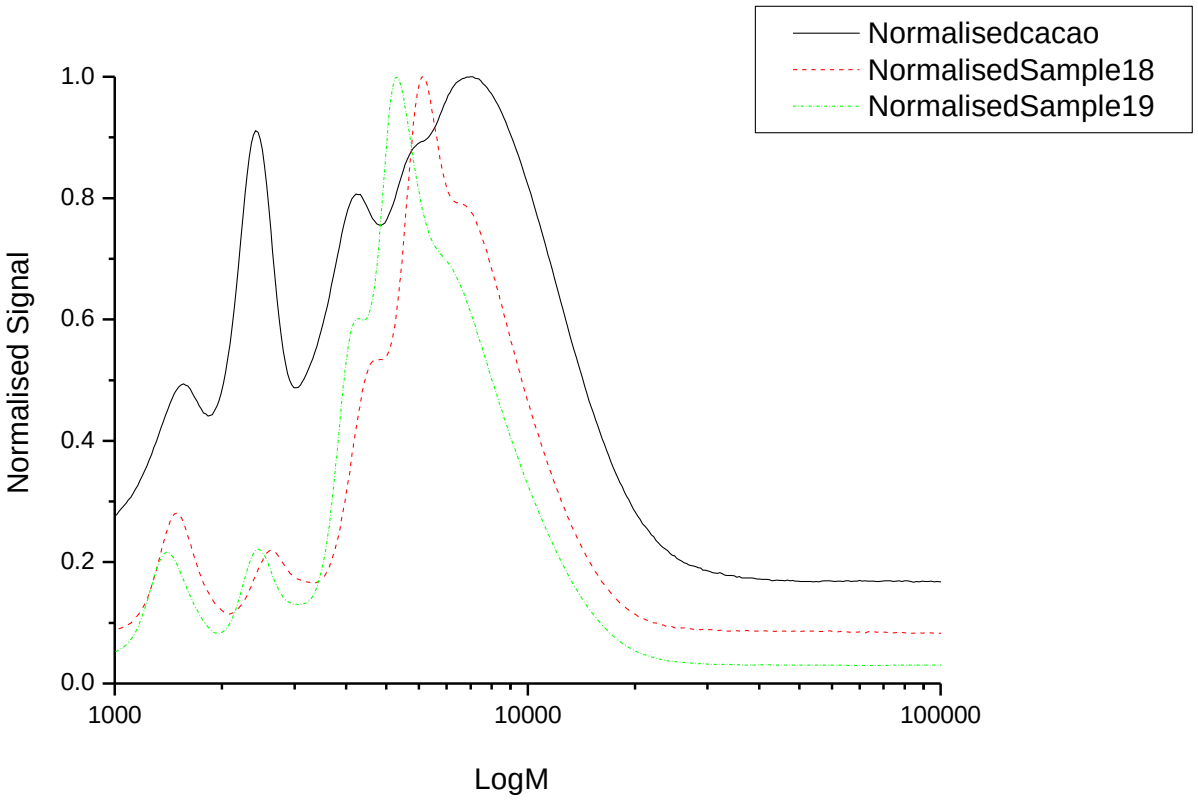


Figure 4.17: Signal response vs. LogM plots obtained by SEC of mimosa tannins, samples 18 and 19 overlaid with cacao tannin, calculated using the PMMA calibration. For sample details refer to Table 3.1. Stationary phase: 2 x PLgel mixed- D columns (300 × 7.5 mm i.d., 5µm) connected in series and protected by a guard column with the same material. Mobile phase: DMAc with 0.03 % (w/v) LiCl and 0.05% (w/v) BHT. Flow rate: 1 mL/min, detection: RI.

Although the absolute mass of the extracts could not be determined this separation still has the added advantage over the NP-LC methods often used to separate oligomeric tannins in that separation is independent of adsorption and thus the molecules are solely separated by their size in solution. The molar masses shown in the table are relative, due the unavailability of higher molar mass tannin standards.

4. 4. References

[1] A. Pizzi, Wood Adhesives Chemistry and Technology, Marcel Dekker, New York, 1983.

[2] R.S. Thompson, D. Jacques, E. Haslam, R.J.N. Tanner, J. Chem. Soc., Perkin Trans. 1 (1972).

[3] E. Haslam, Practical polyphenolics from structure to molecular recognition and physiological action, Cambridge University Press, New York, 1998.

[4] A. Pizzi, Advanced Wood Adhesives Technology CRC Press, New York, 1994.

[5] K.M. Kalili, A. de Villiers, J. Chromatogr. A 1216 (2009) 6274.

[6] M.A. Kelm, J.C. Johnson, R.J. Robbins, J.F. Hammerstone, H.H. Schmitz, J. Agric. Food. Chem. 54 (2006) 1571.

[7] C. Perret, R. Pezet, R. Tabacchi, Phytochem. Anal. 14 (2003) 202.

[8] B. Labarbe, V. Cheynier, F. Brossaud, J.-M. Souquet, M. Moutounet, J. Agric. Food Chem. 47 (1999) 2719.

[9] A. Yanagida, T. Shoji, Y. Shibusawa, J. Biochem. Bioph. Methods 56 (2003) 311.

[10] A. Yanagida, T. Kanda, T. Shoji, M. Ohnishi-Kameyama, T. Nagata, J. Chromatogr. A 855 (1999) 181.

[11] A. Pizzi, D. Thompson, J. Appl. Polym. Sci. 55 (1995) 107.

[12] J.F. Hammerstone, S.A. Lazarus, A.E. Mitchell, R. Rucker, H.H. Schmitz, J. Agric. Food. Chem. 47 (1999) 490.

[13] J. Rigaud, M.T. Escribano-Bailon, J. Prieur, J.-M. Souquet, V. Cheynier, J. Chromatogr. A 654 (1993) 255.

[14] C.S. Ku, S.P. Mun, Wood Sci. Technol. 41 (2007) 235.

[15] A. Yanagida, H. Murao, M. Ohnishi-Kameyama, Y. Yamakawa, A. Shoji, M. Tagashira, T. Kanda, H. Shindo, Y. Shibusawa, J. Chromatogr. A 1143 (2007) 153.

[16] J.A. Kennedy, A.W. Taylor, J. Chromatogr. A 995 (2003) 99.

[17] L.K. Kostanski, D.M. Keller, A.E. Hamielec, J. Biochem. Biophys. Methods 58 (2004) 159.

[18] H. Pasch, A. Pizzi, K. Rode, Polymer 42 (2001) 7531.

Chapter 5

Partial Characterisation of Some Commercial Polymeric Hydrolysable Tannins

5.1. Introduction

Hydrolysable tannins are gallic acid or hexadroxydiphenic acid esters of glucose or other polyols [1,2]. The hydrolysable tannins are further subdivided into simple gallotannins which are mixtures of simple phenols and ellagitannins which are esters of gallic and digallic acids attached to a core sugar which is usually glucose [3]. These tannins have also been shown to be of high molar mass in recent studies [3,4]. The tannins can be composed in various plants and trees such as chestnut, tara, chinese gall, turkey gall, sumac and oak tannin. The monomers that form gallotannins are pentagalloylglucopyranose and the precursor for ellagitannins is trigalloyl-HHDP-glucopyranose [3,5-8]. The different types of hydrolysable tannins structures are shown in **Figure 5.1**. In nature however there exist different oxidation products and polymerised forms of the molecules described. In the analysis of this extract, other higher oxidation structures were isolated, such as castalin/vescalin and vescalagin/castalagin and the structures shown in **Figure 5.2**. Pizzi et al. showed that although these molecules are composed in significant amounts in tannin extracts they are the result of extraction and in situ in the wood the tannins exist as oligomers [3].

The analysis of commercially extracted hydrolysable tannins is not wide-spread compared to their counterparts the condensed tannins. For the commercially available hydrolysable tannins previous studies have focussed on tannic acid which is a mixture low molar mass gallotannins. Analysis of oligomeric hydrolysable tannins has been carried out to a lesser extent and of these chestnut tannin has been the one studied the most [1,3,4,9-11]. The industrial extraction methods alter the composition of the tannin extract [3]. Depending on the solvent composition used there is preferential extraction of certain oligomers and it is well known that solvent extraction removes unattached carbohydrates from the tannin extract. The function of the hydrolysable tannins is affected by structure and this has been investigated in the fields of food science and not so much on the commercial tannin extracts [8,9,12]. Structural characterisation of oligomeric hydrolysable tannins has been carried out by various methods over the years. For molar mass determinations ESI-MS was the popular choice in structural elucidation,

until only recently Pizzi et al. showed the applicability of MALDI-TOF MS as the method of choice in the analysis of hydrolysable tannins [3,4]. This technique was able to show the presence of higher molar mass structures not observed by other techniques. In order to elucidate chemical structures NMR both ^{1}H and ^{13}C have been utilised. However as well known, interpretation of such spectra can get complicated with higher structures and although ^{13}C NMR spectra prove simpler to interpret due to higher resolution, they offer limited information on single molecules.

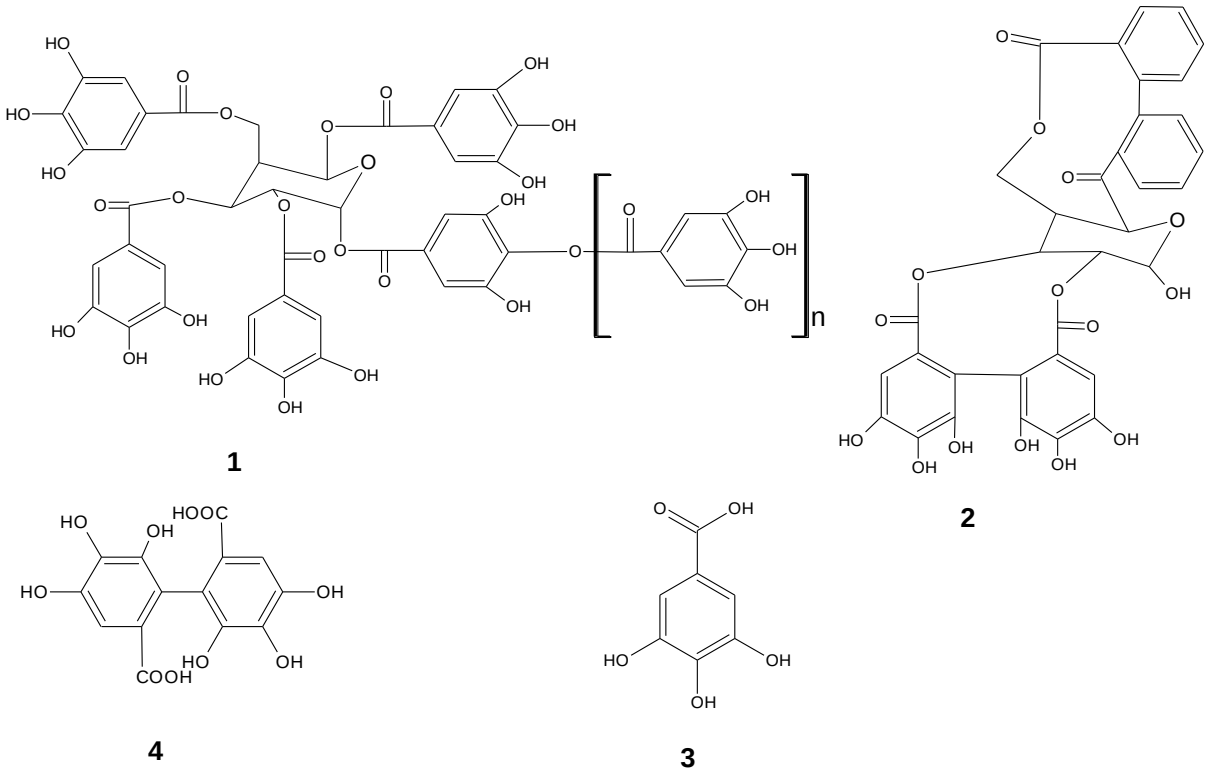


Figure 5.1: Structures of hydrolysable tannins. 1. a gallotannin. 2. an ellagitannin. 3. gallic acid. 4. hexahydroxydiphenic acid (HHDP).

In addition to spectroscopic techniques efforts have been made to achieve chromatographic separation and characterisation. This has been achieved in part for some hydrolysable tannin extracts but as mentioned the study on the commercial oligomeric tannins has been limited in this respect. Normal

phase liquid chromatography can be used to separate oligomers according to the degree of polymerisation [1,7,11,13]. The separation however is by interaction chromatography and thus the separation is influenced not only by the size of the molecule but also its structure [12]. In order to be able to separate the molecules solely by size a separation technique such as size-exclusion chromatography is required. In 'normal' cases the tannin molecules are methylated or acetylated in order to prevent tailing in SEC which is caused by interaction with the styrene-divinylbenzene stationary phase [14,15]. These chemical preparations are tedious and thus the focus has been to find more direct methods to analyse the tannin molecules

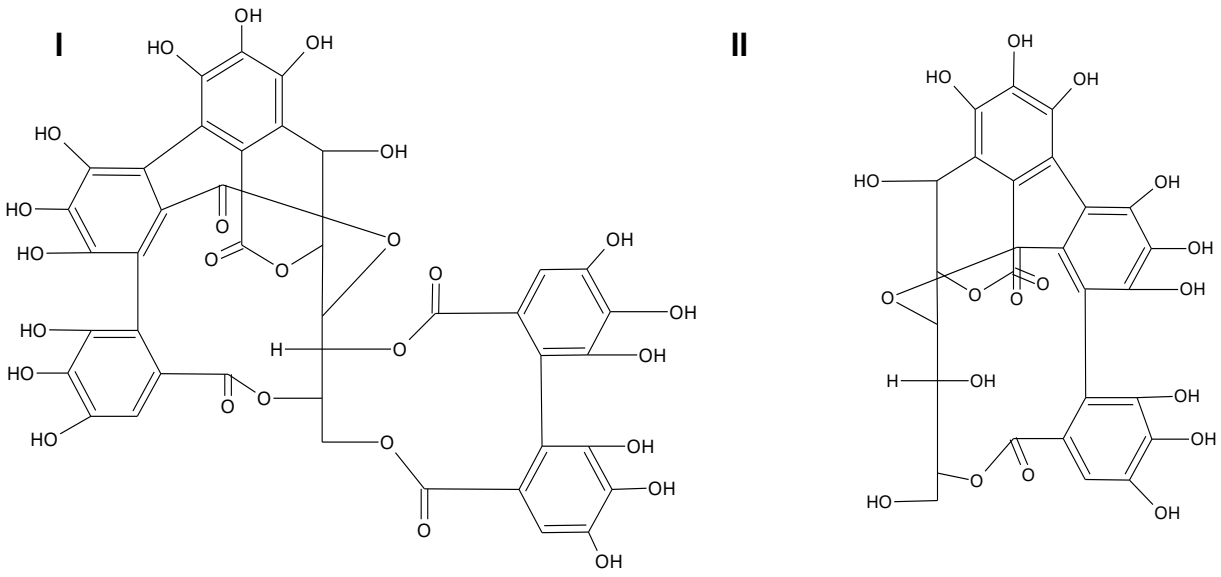


Figure 5.2: Structures of (I) castalagin/vescalagin and (II) castalin/vescalin positional isomers extracted from chestnut tannin [3].

In the sections that follow the analyses of commercial hydrolysable tannins from solvent extracted tara, chestnut and turkey gall are shown. The chemical composition and molar mass distributions of the three hydrolysable tannins are investigated and compared. The tara and turkey gall tannins are known to be gallotannin whereas chestnut is an ellagitannin.

5.2. Experimental

5.2.1. Reagents and materials

All solvents (acetonitrile, methanol, dimethyl acetamide, dimethyl formamide, acetone and acetic acid) were of HPLC grade and purchased from Sigma-Aldrich (Steinheim, Germany). The deionised water was obtained using a Milli-Q water purification system (Millipore, Milford, MA, USA). All the crude samples were filtered through 0.45 µm nylon filter (Membrane solutions Inc.).

5.2.2. Sample preparation

Three types of commercial hydrolysable tannin extracts were used for all the analyses, namely (i) chestnut (*Castanea sativa*) wood tannin, solvent extracted, an ellagitannin; (ii) tara (*Caesalpina spinosa*) tannin, solvent extracted with ethylacetate/ acetone/ water, a gallotannin (iii) turkey gall, the abnormal growth on the stems and branches of certain oaks (*Quercus spp.)* produced by wasps (*Cynips tinctoria*); a gallotannin [3]. All these samples were provided courtesy of Antonio Pizzi.

5.2.2.1. MALDI-TOF analyses

4mg of each tannin sample was dissolved in a mixture of acetone/water 50/50 (% v/v). A solution of 10 mg/mL 2,5-dihydroxybenzoic acid (DHB) was used as matrix. The polymer solution was mixed with the matrix in a ratio of 1:1 and 0.3 µL (10 mg/mL) NaCl was added in order to enhance ion formation. The resulting solutions were placed on the MALDI target and the solvent was allowed to evaporate in air.

5.2.2.2. ^{13}C NMR analyses

A solution of 50/50 (% v/v) acetone/water was prepared, followed by preparation of a mixture of d_6-acetone/D_2O in the same ratio. These two solutions were combined and 1 mL this mixture was used to dissolve 300 mg of each sample. In order to facilitate dissolution the samples were placed in an ultrasonic bath for 5 minutes. This allowed for further dissolution; however, at this relatively high concentration the tannins were only partially soluble. The resulting sample solutions were then filtered using 0.45μm nylon filter then transferred to an NMR tube to be analysed.

5.2.3. Instrumentation

5.2.3.1. MALDI –TOF CID analyses

MALDI-TOF analyses were performed on a AximaTOF^2 spectrometer (Shimadzu Biotech, Manchester, UK), equipped with a nitrogen laser (337 nm),the pulsed extraction ion source accelerated the ions to a kinetic energy of 20 keV.All analyses were carried out in the linear positive mode. Calibration was done using 1450 Da PEG standard. The instrument is equipped with an ion gate in order to select precursor ions and a collision cell. Argon was used as the collision gas and the precursor ions were separated by approximately 4 Da mass window. The pressure in the collision cell was 8 x 10^{-6} mbar. The ions were accelerated to a kinetic energy of 20 keV. All the data was obtained in the positive reflectron mode.

5.2.3.2. ^{13}C NMR analyses

The ^{13}C NMR (300 or 400 MHz) analyses were recorded on a Varian 300 VNMRS or Varian Unity Inova 400 instrument. Both instruments are equipped with two channels and probes of varying sizes. TMS

was used as the internal standard. The samples were run over the weekend whereas the standards were run overnight. The decoupling mode with NOE was used.

5.2.3.3. SEC analyses

Analyses were carried out on an instrument equipped with an isocratic pump (Shimadzu, Japan), autosampler, column oven, dual wavelength detector and a refractive index detector all from Waters. The data acquisition was performed by Millennium software. The system was calibrated with polymethyl methacrylate polymers from easi-vials, purchased from Polymer Labs (Amherst, MA, USA) were used in calibration. All the sample were filtered with 0.45 µm filter (Membrane Solutions Inc)

5.2.3.4. NP-HPLC analyses

The normal phase liquid chromatography analyses were carried on an Agilent 1200 series HPLC instrument equipped with vacuum degasser, a quaternary pump, autosampler, column oven and variable wavelength detector (Agilent Technologies, CA, USA). The data was acquired using the Chemstation software.

5.2.4. Chromatographic methods

5.2.4.1. Size exclusion chromatography (SEC) analyses

The separation of the oligomeric hydrolysable tannins were performed on a two PLgel mixed- D columns (300 × 7.5 mm i.d., 5 µm d_p) connected in series and protected by a guard column with the

same material purchased from Polymer Labs (Amherst, MA, USA). The mobile phase used was DMAc with 0.03 % (w/v) LiCl and 0.05 % (w/v) BHT at a flow rate of 1 mL/min. The column temperature was kept at 40°C. The samples were monitored using UV at 280 and 320 nm as well as RI detection. All the samples were dissolved in pure DMAc solvent before being introduced into the column. This method was adapted from Kennedy et al. [16]. Calibrations were performed using polymethylmethracrylate polymers of molar mass with 850, 2810, 4900, 10290, 30530, and 60150 Da.

5.2.4.2. Normal phase high performance liquid chromatography (NP-HPLC) analyses

The normal phase separations were carried out on a Nucleosil unmodified silica column (250 x 4.6 mm i.d., 5 µm, 100 Å) Machery Nagel (Separations, Randburg, SA). The mobile phases used were A: n-Hexane and B: 3/1 methanol/THF with 0.25 % acetic acid. Isocratic runs were first carried out, however this did not provide sufficient resolution therefore various gradients were introduced. The mobile phase composition was 58 % B for 35 minutes. A UV detector was used at 280 nm. Needle wash consisting of 99/1 % (v/v) ACN/HOAc was used to rinse the needle after each injection. The samples were prepared in solvent B (10 mg/mL) and 10 µL of this was introduced into the column.

5.3. Results and discussion

5.3.1. Analysis of molar masses by MALDI-TOF MS

The hydrolysable tannins were analysed by MALDI-TOF MS and as illustrated with the condensed tannins (Chapter 3) this technique is successful in determining both the molar mass and chemical composition. The tara and turkey gall tannins are both gallotannins with the polygallic form [3]. Pizzi et al. utilised MALDI-TOF MS to investigate the structure and degree of polymerisation of the extracts [3].

Figure 5.4 and 5.5 show the MALDI-TOF spectra of the hydrolysable tannins from tara and turkey gall, respectively. As expected a similar repeat unit of 152 Da is observed for both tannins and the structure is shown in **Figure 5.3**. The structural assignments were done based on the work of Pizzi et al.[3].

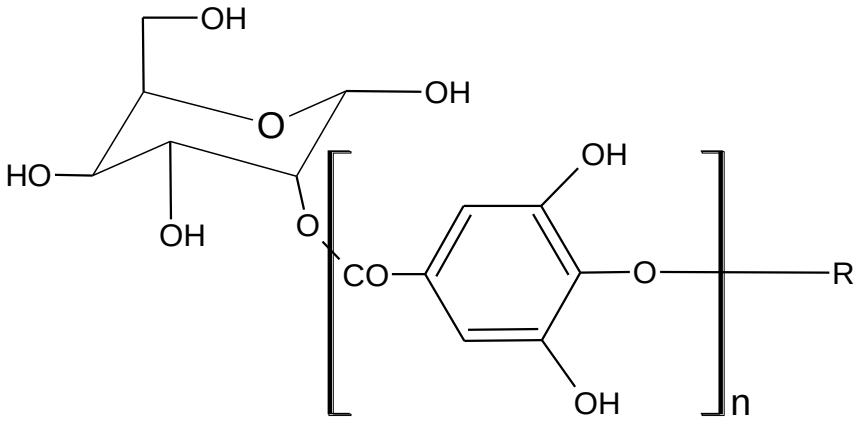


Figure 5.3: The proposed structure of the gallotannin extracted from turkey gall tannin.

5.3.1.1. The gallotannins: Tara and turkey gall tannin

The peaks at 521, 673, 825, 978, 1130, 1282 and 1435 Da in the tara tannin represent the major series and are separated by 152 Da which belong to a gallic acid unit. The oligomer at 520 Da is formed by a single ellagic unit, gallic acid unit and an additional 44 Da from a $–COO^-$ group. Thus the following equation was used to construct **Table 5.1**, $M+Na^+ = 23(Na^+) + 302\text{-}1H$ (ellagic unit) + 152 (gallic unit) + 44 (COO^-). This means that the oligomers are detected up to heptamers.

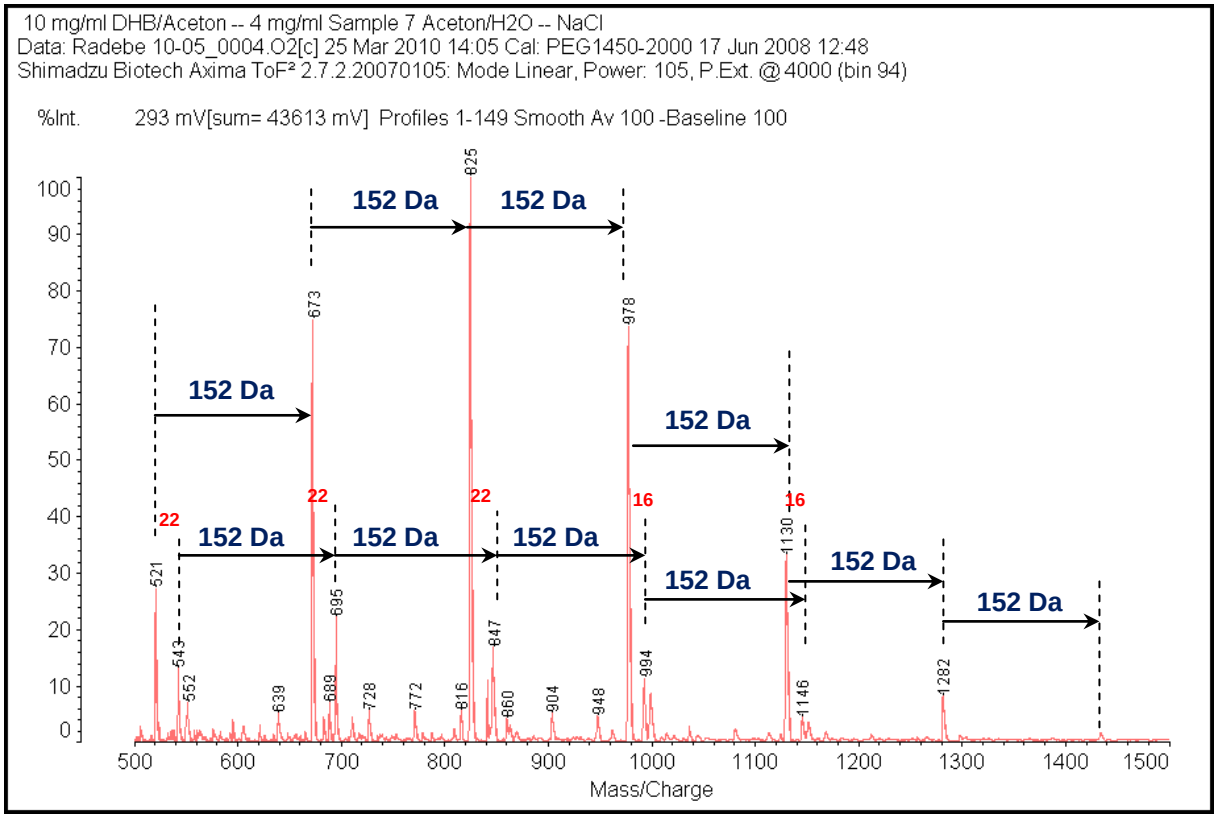


Figure 5.4: MALDI-TOF spectrum of solvent extracted tara tannin, showing a repeat unit of 152 Da. Mass range: 500-1600 Da.

A series of peaks with lower intensity appears adjacent to the major series and its peaks also have a repeat unit of 152 Da. However, this minor series is 16 Da from the major series and thus is made up of different but related structure. These peaks represent the oligomer preceding it minus an –OH. For example, the peak at 1146 Da represents the pentamer with an additional -OH group. Therefore the structure of the tara tannin is a relatively simple one. Although this extract consists of mainly gallic acid oligomers some ellagic units are present as well, making up the major oligomeric series.

Table 5.1: Dominant oligomer MALDI-TOF peaks and description of structures present in the solvent extracted tara tannin.

Calculated $M+Na^+$	**Experimental** $M+ Na^+$	**Description** $M+ Na^+$
520	521	$-COO^-$ + Ellagic acid+ 1 gallic acid
672	673	$-COO^-$ + Ellagic acid+ 2 gallic acid
824	825	$-COO^-$ + Ellagic acid+ 3 gallic acid
976	978	$-COO^-$ + Ellagic acid+ 4 gallic acid
1128	1130	$-COO^-$ + Ellagic acid+ 5 gallic acid
1280	1282	$-COO^-$ + Ellagic acid+ 6 gallic acid
1432	1435	$-COO^-$ + Ellagic acid+ 7 gallic acid

A look at the turkey gall tannin also shows simple polygallic structure. The basic structure of the turkey gall is similar to that of the tara tannin but does not include any ellagic acid groups in its basic structure. The 508 Da peak observed by Pizzi et al. is observed in very low intensity in the spectrum shown in **Figure 5.5** [3]. The first series of peaks is observed from the 508 Da peak which is representative of a sodium adduct of two gallic acid units attached to glucose (**Figure 5.3**). The 661 Da peak is 152 Da separated from the digalloyl glucose, which means it is a trimer and thus from this spectrum the oligomers are detected only from trimers. Although the two tannins have the same repeat unit, the basic unit is definitely different to that of the tara tannin. In this case the gallic acid units are added to a single glucose molecule. The peaks that belong to this extract are shown in **Table 5.2**, $M+Na^+ = 23(Na^+) + 152$ (gallic unit).

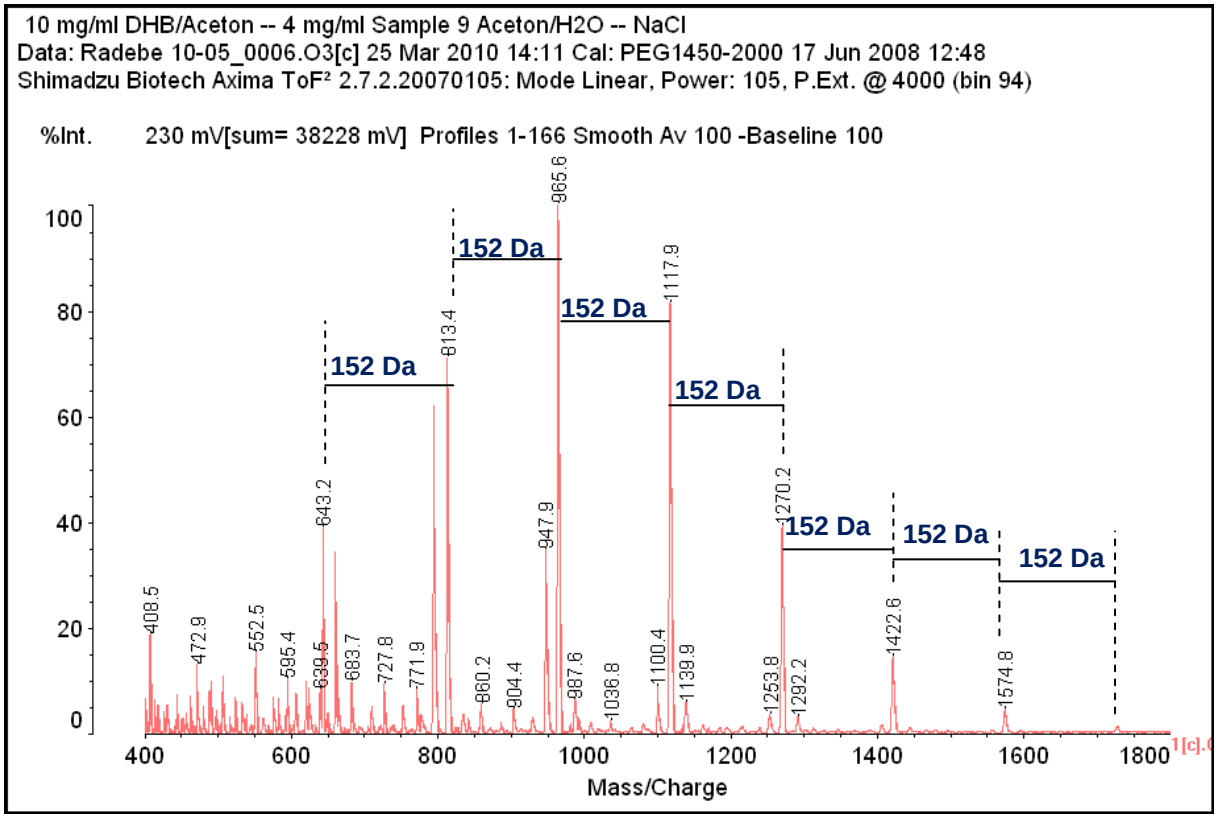


Figure 5.5: MALDI-TOF spectrum of solvent extracted turkey gall tannin with repeat unit of 152 Da indicated. Mass range: 400-1900 Da.

The dominant oligomer peak at 965.9 Da is believed to represent pentagalloyl glucose, and no higher oligomers of this type are detected. As suggested by Pizzi et al. it may due to the fact that the higher oligomers are destroyed during extraction from the wood [3].

5.3.1.2. Chestnut tannin

The structures of natural water-extracted chestnut tannin and the chestnut tannin treated with 3.5 % ammonium sulphite for use in leather tanning obtained by MALDI-TOF has already been reported [3,4]. Analysis of both extracts proved the oligomeric nature of the extract in situ in the wood, the tannin was shown to be composed of pentagalloylglucose (PGG) oligomers with up to 16 units [3,4]. In the water-

extracted chestnut tannin, the PGG oligomers were detected up to trimers whereas in the sulphited sample glucose oligomers stripped of its galloyl units was detected [3] . The linear mode spectrum of solvent extracted chestnut tannin is shown in **Figure 5.7**; this spectrum is markedly different from the previously reported spectra for this tannin. The peak assignments were made after Pizzi et al [3].

Table 5.2: Dominant oligomer peak series and description of structures as observed from the MALDI-TOF spectrum of solvent extracted turkey gall tannin.

Calculated $M+Na^+$	**Experimental** $M+Na^+$	**Description** $M+Na^+$
508	509	Digalloyl glucose
659	661	Glucose + 3 gallic acid
811	813	Glucose + 4 gallic acid
963	966	Glucose + 5 gallic acid
1115	1118	Glucose + 6 gallic acid
1267	1270	Glucose + 7 gallic acid
1419	1422	Glucose + 8 gallic acid
1571	1575	Glucose + 9 gallic acid

The mode of extraction clearly affects the composition of this tannin significantly. The repeat unit in this case is 162 Da which cannot be explained by the theories previously presented for the structures present in this extract, thus this warrants an in depth discussion. The repeat unit of 162 Da is similar to the one observed for the chinese gall tannin. The explanation for this distribution was given as oligomers formed by esters of gallic and digallic units attached to glucose [3]. By calculation only it

cannot be determined whether two gallic acid units attached to each other or independently on the pyranose ring of the glucose. The proposed structure of the oligomer is shown in **Figure 5.6**.

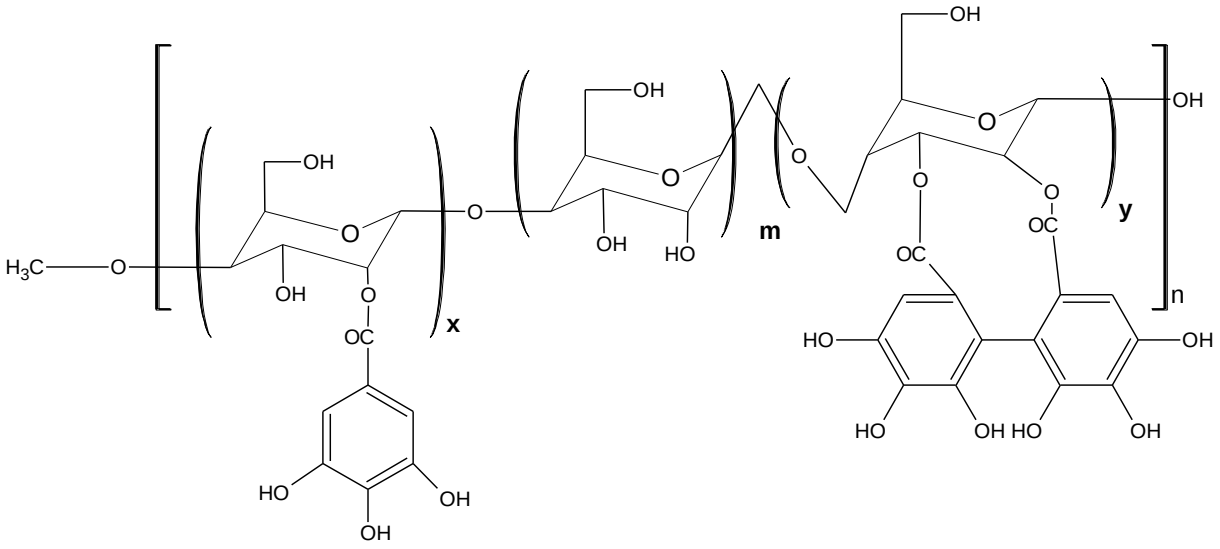


Figure 5.6: Representation of the possible polymer structure present in chestnut tannin. With the values of x, m, and y differing according to chemical structure, n represents the chain length.

The most abundant oligomer distribution is observed with the series of peaks beginning at 508 Da. This peak is known to be a digalloylglucose monomer (digalloylglucose = 23 (Na^+) + 152 x 2 + 180 = 507 Da), the series with 162 Da is completed at 1972 Da. The 162 Da increment represents an additional glucose unit being added to the oligomer structure. The presence of the digalloylglucose monomer supports this interpretation. A list of peaks and assignments is shown in **Table 5.3**. This means that the structures observed during solvent extraction resemble those in the sulphited tannin extract. However in this extract, the dominant series is composed of glucose units that are attached to two galloyl units instead of one. An alternative interpretation was also given for this major series for the chinese gall tannin, the 162 Da repeat unit might just be a glucose molecule that has lost a hydroxyl group. In this discussion the latter explanation is taken as more likely since there are peaks that are 16 Da from each of the dominant oligomer peaks, forming thus a repeat unit of 178 Da. This then supports the theory that the chestnut tannin solvent extract is composed of glucose oligomers attached to 2 galloyl units. The

distribution which is 16 Da from the major series is present up to high molar masses this means that the main composition of this extract is long (up to 11 units) chains of glucose units each attached to two galloyl units. This extract is composed of multiple distributions; however these seem to not form part of the higher molar mass oligomers. The peak observed at 438 Da can be assigned to a single galloyl unit attached to two stripped glucose units. This forms the second distribution of peaks 438, 727, 879 and 1036, which have a mass increment of 152 Da indicative a galloyl unit. The difference of 388 Da observed between the 1972 and 2310 peaks indicates the presence of an ellagic acid unit attached to a –O-C-O- group added to the oligomer chain, (2310 Da = 23 (Na^+) + 2 x 152 + 180 + 162 x 9 + 302 + 44). This assignment may also be made according to Pizzi et al as 2 additional galloyl units being added to the stripped glucose skeleton structure, the former is less likely however it is possible that these structures may be present since the chestnut tannin is a known ellagitannin [3].

The analysis of the chestnut tannin by reflectron mode shows a different spectrum from the analysis in the linear mode, it shows less peaks in the low molar mass region. This is due to the different methods in which the ions are detected. In the reflectron mode only stable ions are detected. The mass increment of the dominant series is changed to 152 Da, as shown in **Figure 5.8** but in addition to this there are peaks that show the 162 Da mass increments as observed in the linear mode spectrum. The reflectron mode spectrum obtained more closely resembles the one obtained for the bisulphited tannin in the linear mode [3]. The structure that was elucidated using the linear mode thus confirms the presence of higher PGG oligomers in situ the wood.

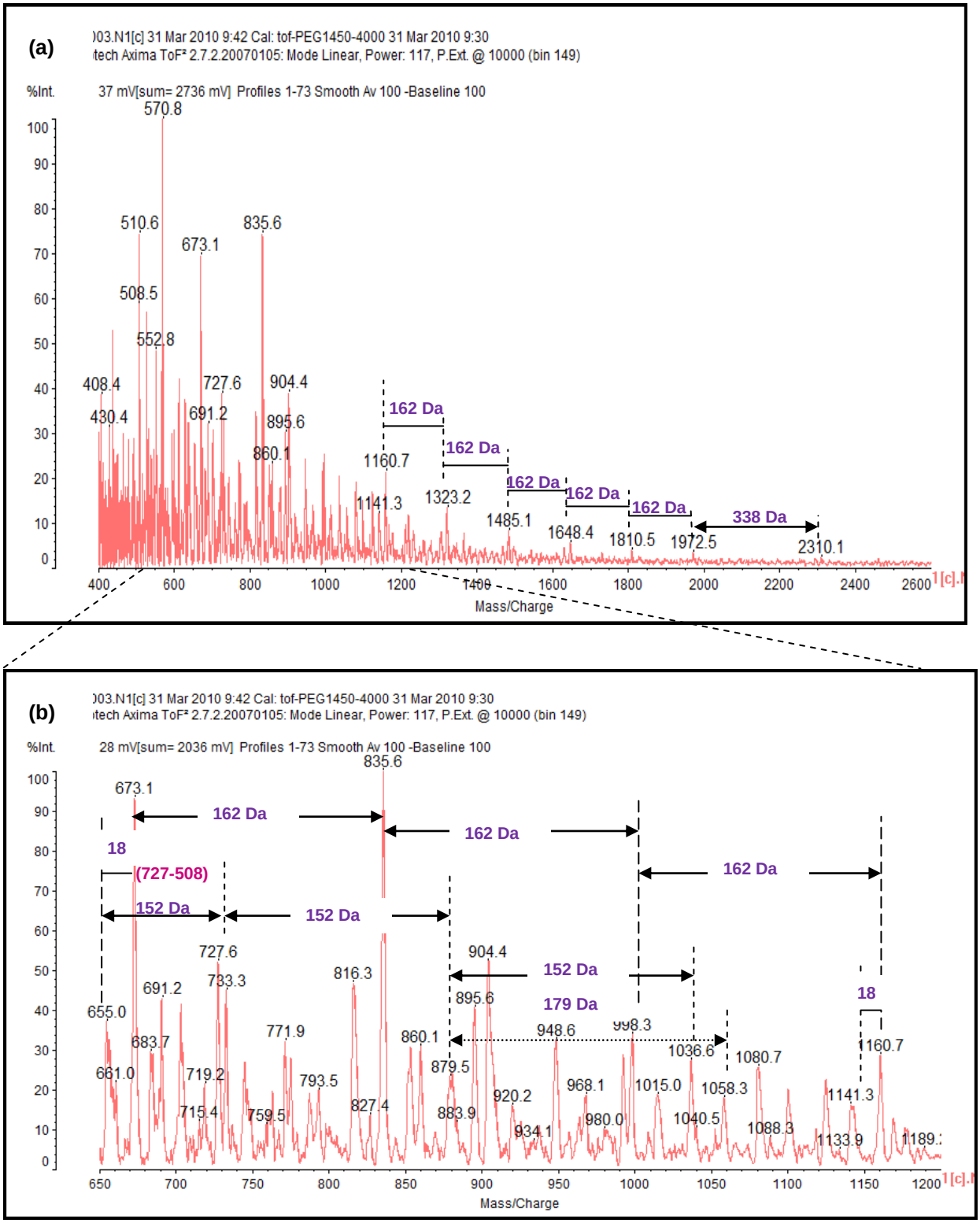


Figure 5.7: MALDI-TOF spectrum of solvent extracted chestnut tannin (linear mode). Mass range: (a) 400-2600 Da (b) 650-1200 Da. The peak assignments are indicated and the details shown in Table 5.3.

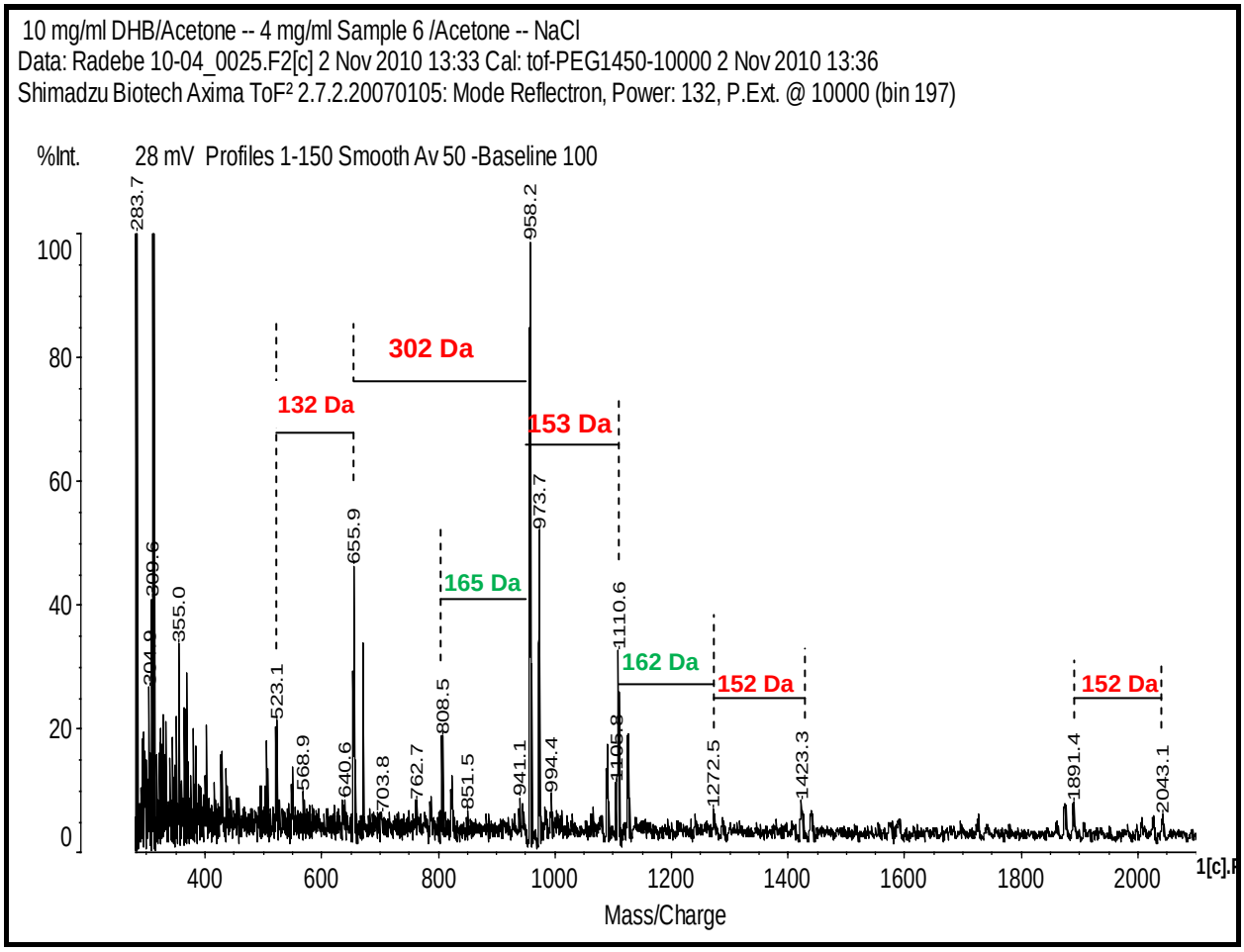


Figure 5.8: MALDI-TOF spectrum of solvent extracted chestnut tannin (reflectron mode). Mass range: 300-2200 Da. The mass increments of 152 Da and 162 Da are indicated.

Solvent extraction of the chestnut tannin reveals that it is composed of various types of structures. In the water-extracted tannin PGG oligomers up to trimers were detected and in this extract part of the peak observed at 965 Da could belong to the PGG monomer. The presence of glucose oligomers in the solvent extracted sample shows that multiple types of structures are contained in the chestnut tannin. The structures whereby the PGG forms the nucleus and galloyl units are added to form long oligomers, or where PGG monomers combine to form longer chains are all possibilities. A third type of structure is when an ellagic acid unit is added instead. Long chains of ellagic acid units have been observed in the chestnut tannin [10]. The glucose oligomers indicate the structure where the galloyl units are attached around polymerised glucose monomers (**Figure 5.14**).

Table 5.3: Observed and calculated oligomer peaks from the linear mode MALDI-TOF spectrum of solvent extracted chestnut tannin.

Experimental ($M+Na^+$)	Calculated ($M+Na^+$)	Description ($M+Na^+$)	Oligomer
438	439	1 Galloyl unit + 2 stripped glucose	Dimer
508*	507	Digalloyl glucose, monomer	Monomer
571*	571	1 Galloyl unit + 3 stripped glucose	Trimer
657*	659	Trigalloyl glucose, monomer	Monomer
672*	669	Digalloyl glucose + dehydrated glucose	Dimer
835*	835	1 Galloyl unit + 5 stripped glucose	Pentamer
968	967	1 Galloyl unit + 6 stripped glucose	Hexamer
998	993	Digalloyl glucose + 3 dehydrated glucose	Tetramer
1161*	1155	Digalloyl glucose + 4 dehydrated glucose	Pentamer
1233	1232	1 Galloyl unit + 8 stripped glucose	Octamer
1323*	1317	Digalloyl glucose + 5 Dehydrated glucose	Hexamer
1364	1364	1 Galloyl unit + 9 stripped glucose	Nonamer
1485*	1479	Digalloyl glucose + 6 dehydrated glucose	Heptamer
1630	1627	1 Galloyl unit + 11 stripped glucose	Decamer
1648*	1641	Digalloyl glucose + 7 dehydrated glucose	Octamer
1764	1759	1 Galloyl glucose + 12 stripped glucose	Dodecamer
1810*	1803	Digalloyl glucose + 8 dehydrated glucose	Nonamer
1972*	1965	Digalloyl glucose + 9 dehydrated glucose	Decamer
2155	2155	1 Galloyl unit + 15 stripped glucose	Pentadecamer
2310*	2307	2 Galloyl glucose + 15 stripped glucose	Decamer
2462*	2269	Trigalloyl glucose + 9 dehydrated glucose	Nonamer

dominant oligomer peaks

5.3.2. Oligomer sequence determination by MALDI-TOF MS CID

5.3.2.1. Tara tannin

The tara tannin has been analysed by MALDI-TOF and was shown to have a simple polygallic structure (**Figure 5.10**). The MALDI-TOF spectrum of the bulk sample is shown in the previous section (**Figure 5.4**) and the major repeat unit observed is 152 Da which represents a gallic acid unit as previously discussed. The structure assignment shown in **Figure 5.10** is made according to information available for this tannin however; the monomer sequence represented requires confirmation.

For further investigation of the oligomer structures presented in the tara tannin the precursor ions with m/z 824, 839 and 976 were selected for fragmentation experiments. The first to be considered will be the m/z 824 trimer, $M+Na^+$ = 23(Na^+) + 302-1H (ellagic acid) + 3 x 152 (gallic acid) + 44 (COO^-) shown in **Figure 5.10**. The MALDI-CID spectrum for this oligomer is presented in **Figure 5.9**; the mass loss incurred by the precursor ion and the product ions are also indicated. The fragmentation pattern of the oligomer is relatively simple and the dominant ions appear at m/z 674, 502, 328 and 154 (**Figure 5.9**). This series is formed from loss of a gallic acid unit (152 Da) attached to additional –O- group which results in a mass of 170 Da. The loss of a gallic acid unit is also observed as the second most dominant series and for this oligomer fragmentation occurs three times, representing the loss of the 3 gallic acid units that form the oligomer structure. Once formed, the fragment ions may lose an –O- or C=O group, shown by mass losses of 16 and 28 Da, respectively. These types of fragments show that the galloyl ester bond is the most labile, and thus the fragmentation may occur before or after the carbonyl group of the ester bond. However, as observed in the spectrum, there are no fragment ions resulting from the loss of the –O- and C=O groups simultaneously.

The loss of the 3 gallic acid units as previously noted results in the formation of the m/z 350 fragment ion (**Table 5.4**) which is a proton adduct of ellagic acid unit still attached to the –COO (44) group. Another fragment is observed at m/z 301 which can then be assigned to a proton adduct of the a single ellagic acid unit. The proposed structure arising from the ellagic acid unit at the terminal end of the molecule is supported by the fragmentation pattern.

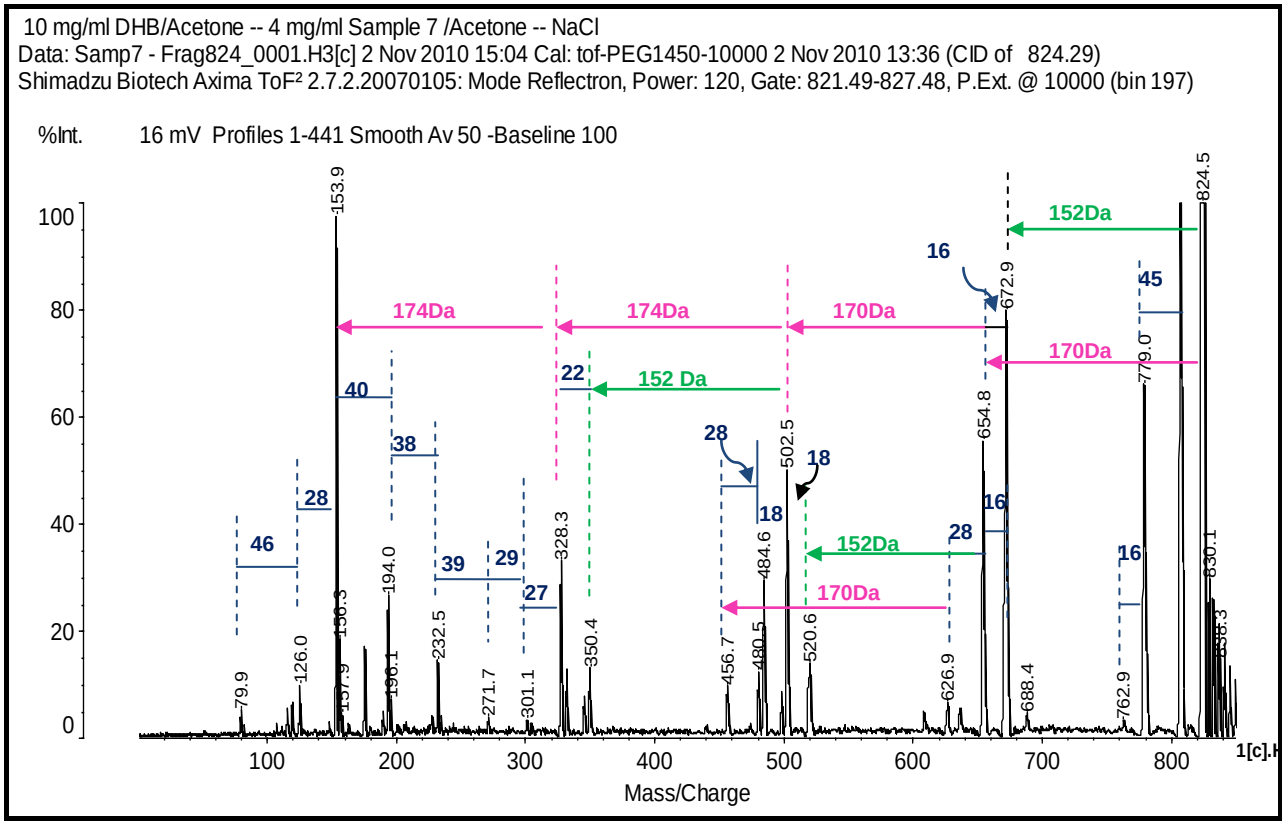


Figure 5.9: MALDI-TOF CID spectrum of the 824.5 Da trimer from tara tannin. The fragmentation pattern is illustrated by indicating the mass loss incurred from the precursor ion.

The precursor ion (m/z 824) additionally looses a mass of 45 Da arising from the –COO to form the m/z 779 ion. This confirms the presence of the postulated 44 Da mass, and the presence of the m/z 350 ion indicates that this group is attached to ellagic acid unit on the oligomer chain as there are no fragments observed that indicate its attachment to the gallic acid units (**Table 5.4**). Furthermore, in the structure

that has been proposed the only likely position is on the ellagic acid unit. An arbitrary position was selected for the location of this group.

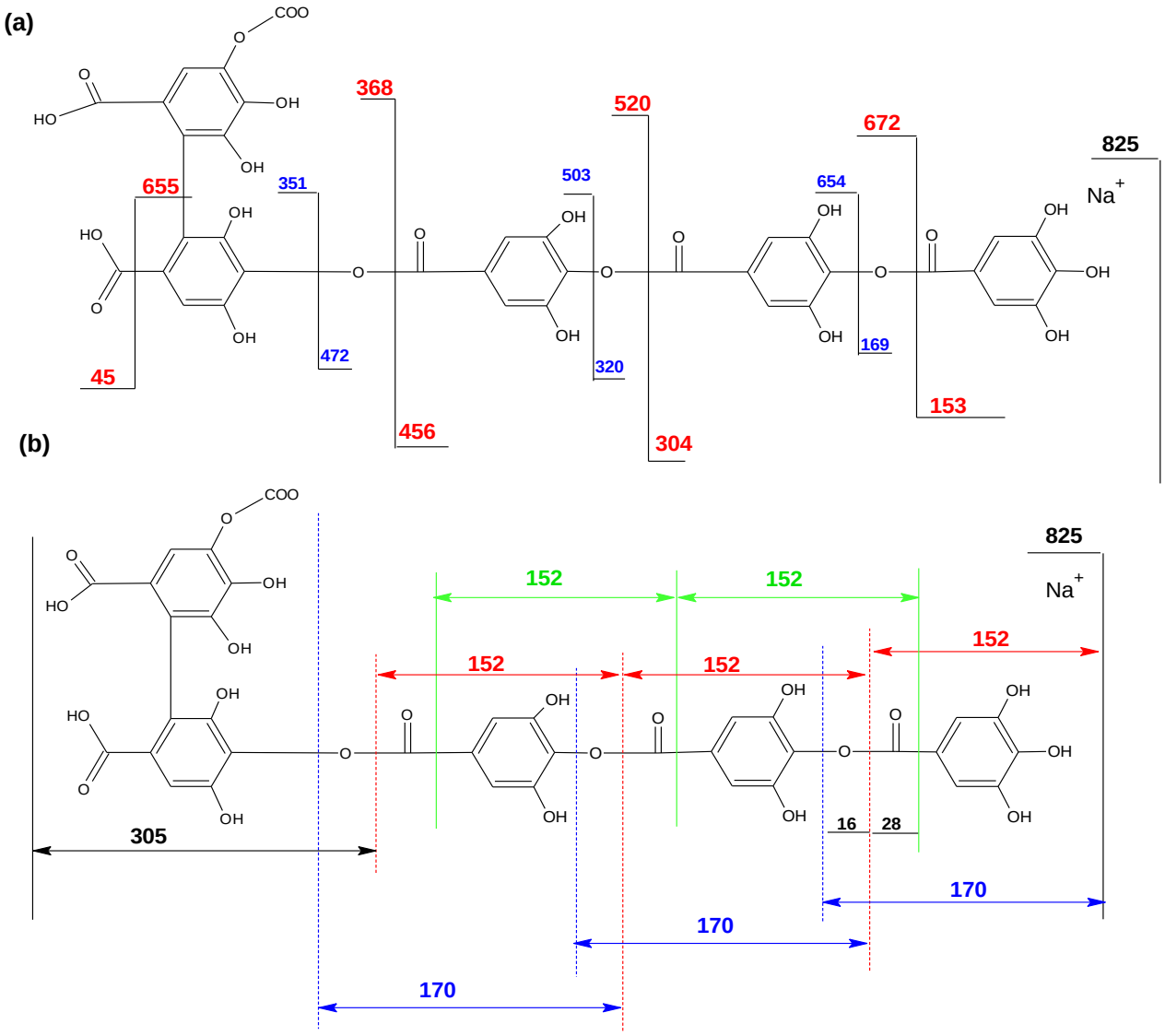


Figure 5.10: Proposed structures and fragmentation pattern for the 824.5 Da trimer observed in the MALDI-TOF spectrum of the tara tannin, the sodium adduct masses are indicated in brackets. (a) Shows some of the possible fragment masses and positions (b) indicates the structures that form the respective repeat units.

Further investigations were conducted on the m/z 839.9 precursor ion. This oligomer is related to the preciously determined m/z 824 structure, by an additional 16 Da mass. From the bulk spectrum of the chestnut tannin the m/z 839.9 ion was attributed to the sodium adduct of trigalloyl diglucose trimer (**Figure 5.12**). This form of structure was suggested due to the polygallic form of the oligomers present in the tara tannin extract, and the MALDI-CID of the m/z 824 ion further enforced this hypothesis. However, when the MALDI-CID spectra of the two ions are compared they show different fragment masses and the only comparable peaks are the m/z 821 and 154 ions. The MALDI-CID spectrum of the m/z 839 ion shows mass losses of 152, 188, 272 and 288 Da, which are very different to what was observed for the m/z 824 ion, that showed a simple fragmentation pattern. The presence of the m/z 154 and 821 fragment ions shows that the structures mentioned are somehow related but the presence of the postulated presence of the glucose rings seems to alter the manner in which the oligomer degrades.

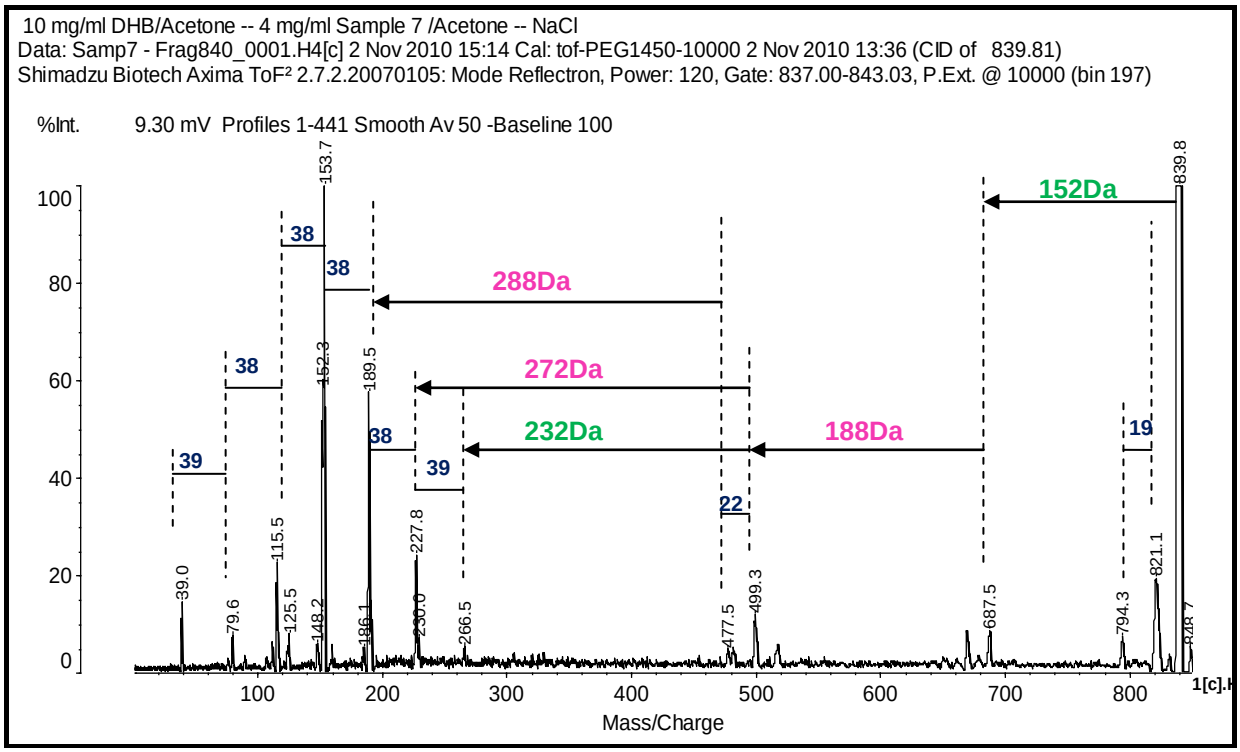


Figure 5.11: MALDI-TOF CID spectrum of the 839.8 Da trimer from tara tannin. The fragmentation pattern is illustrated by indicating the mass loss incurred from the precursor ion.

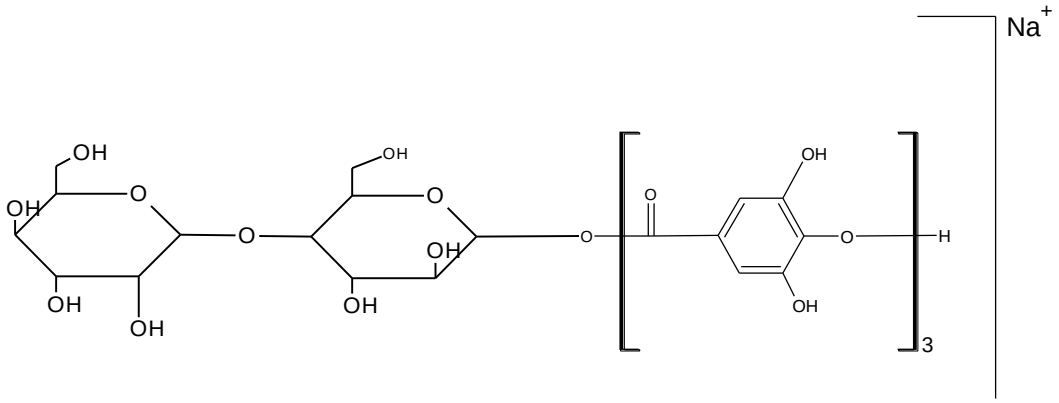
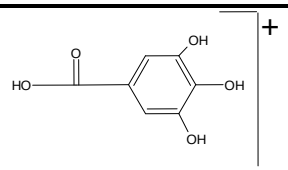
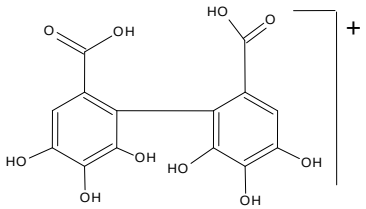
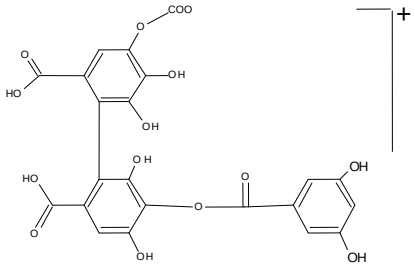
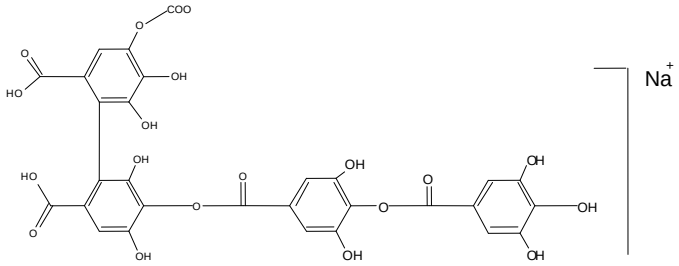
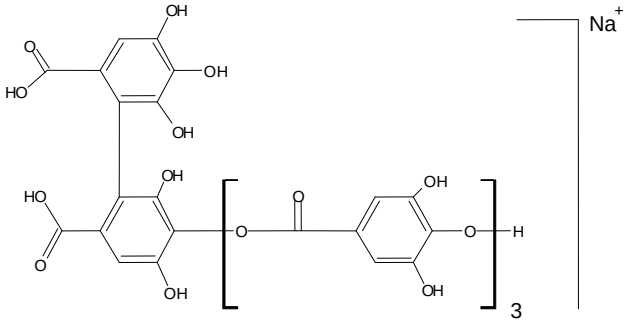


Figure 5.12: Proposed structure for the m/z 839 trimer as postulated from the bulk sample MALDI-TOF analysis.

The repeat units that are observed in the MALDI-CID spectra are not consistent with the proposed structure. As noted for the m/z 824 ion, galloyl esters are very labile and if they were present as indicated in **Figure 5.12**, then the mass losses would be 152 and 170 Da as previously observed. A more detailed look at the MALDI-CID spectrum reveals more detail on the actual arrangement of the gallic acid units in the oligomer structure. The fragment ion at m/z 821 is formed by a mass loss of 18 Da which could be due to a H_2O molecule removed from the structure and the most likely position is on the glucose ring. Further fragmentation occurs via internal fragmentations of the glycosidic linkage represented in **Figure 5.12**, the presence of these internal fragments are by a mass loss of 134 Da from the m/z 821 ion. These types of fragmentations are commonly observed during MALDI-CID of carbohydrates [17]. The internal fragments result from cleavage of the glucose ring in front of the cyclic –O- group. The gallic acid units may be located at any of the –OH groups present in the glucose structure, however, due to the presence of the internal fragments from the glucose molecules, exact locations of these groups may be determined. In this case, one of the gallic acid units must be located at the terminus of the glucose unit; this is shown by the presence of the mass loss of 288 Da observed in the spectrum. This mass can only be formed by the loss of a gallic acid unit attached to a 134 Da fragment from the glucose monomer. Further proof of the presence of the gallic acid unit at the terminal end of the molecule is the presence of the mass loss of 152 Da incurred by the precursor ion.

Table 5.4: Structures of selected fragment ions observed in the MALDI-CID spectrum of 824 Da trimer from the tara tannin.

m/z	Fragment ion
152	
301	
502	
673	
779	

The m/z 665 ion looses a further 188 Da mass, which is a result of –C=O group attached directly to the glucose structure. The structure elucidated after MALDI-CID is shown in **Figure 5.14**, the presence of mass losses 288 and 272 Da can only be formed by the structure as shown. The presence of ellagic acid on the oligomer structure would present with differing fragment ions. Some of the major fragment ions are shown in **Figure 5.15**. The m/z 154 fragment ion structure is shown in **Table 5.4**.

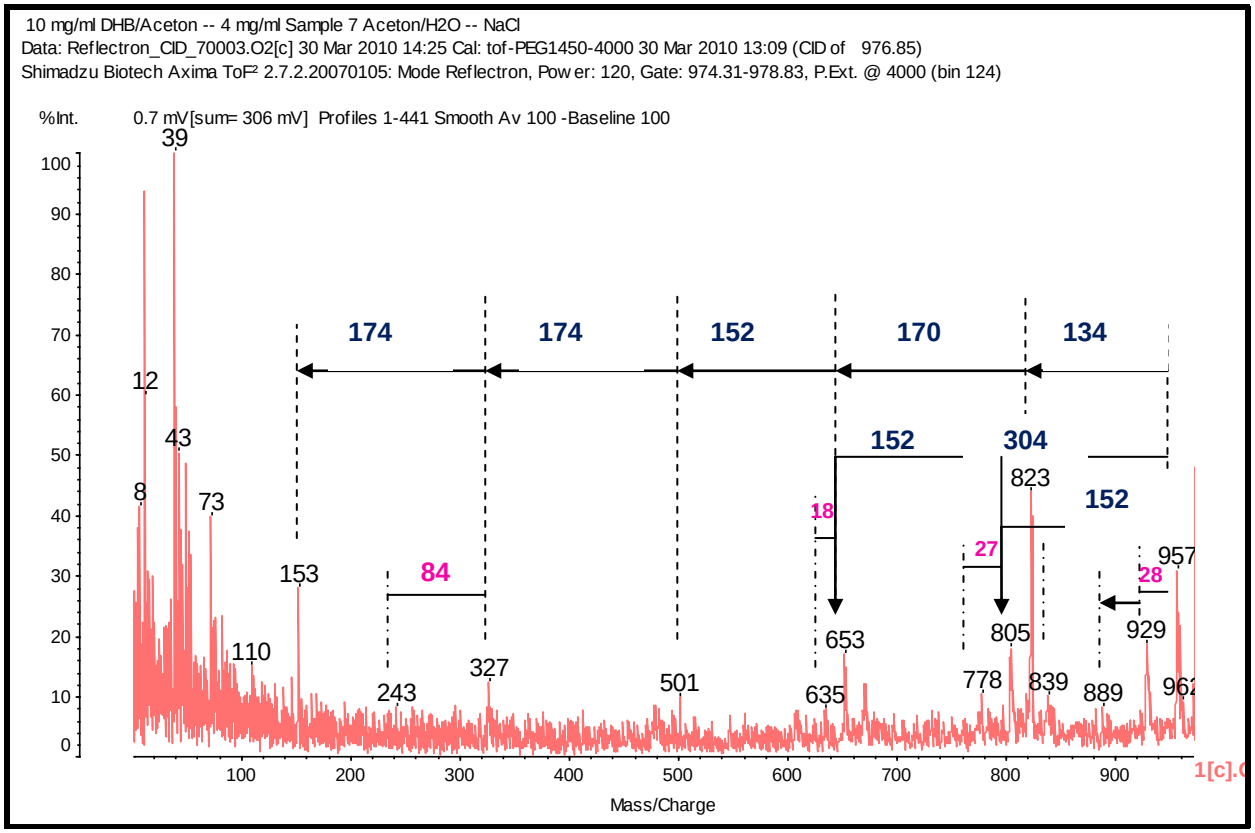


Figure 5.13: MALDI-TOF CID spectrum of the 976 Da trimer from tara tannin. The fragmentation pattern is illustrated by indicating the mass loss incurred from the precursor ion.

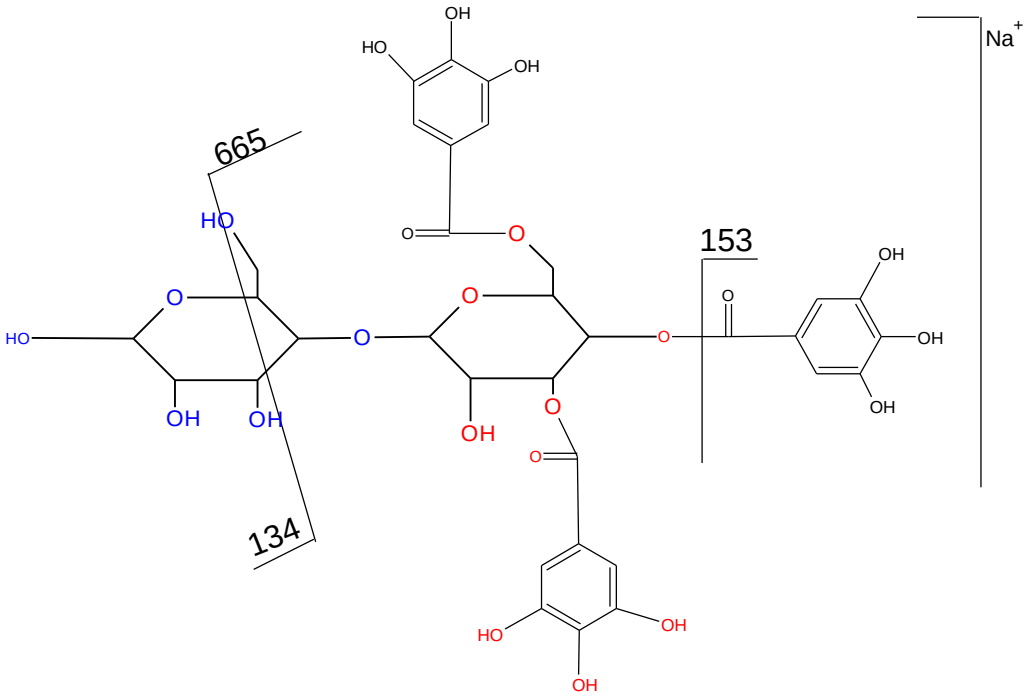


Figure 5.14: The structure of the m/z 839.8 trimer of the tara tannin elucidated by MALDI-CID. Some of the major fragment positions and masses are also indicated.

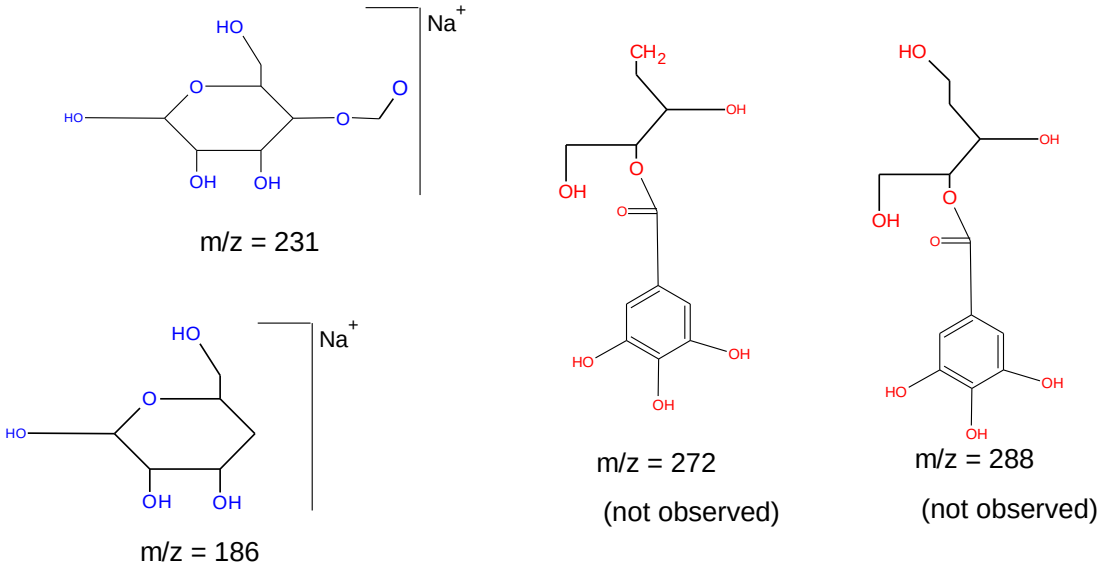


Figure 5.15: Major fragment ions observed in the MALDI-CID of the tara tannin m/z 839.8 trimer. The m/z 288 and 272 are observed as fragments but the actual ions are not detected. The red and blue were used to distinguish between the two glucose rings shown in Figure 5.14.

Further investigations were carried out on the precursor ion with m/z 976 which was assigned to a gallic acid tetramer attached to a single ellagic acid unit and –COO group (**Figure 5.12**). This mass is related to the preciously discussed m/z 824 trimer in that it has an additional gallic acid unit added to the structure. The most important fragments are observed as 823, 778, 653, 501, 327 and 153 which are also present in the MALDI-CID of the m/z 824 trimer. The MALDI-CID spectrum for the m/z 976 precursor ion is shown in **Figure 5.13**. As indicated in this figure the m/z 653 ion results from the loss of two gallic acid units, however it can also be formed by the loss of the ellagic acid, the latter explanation being more likely. The mass loss of 304 Da is not observed for the lower oligomers analysed previously. This ellagic acid unit is only observed once in this fragment ion, which means that the suggested structure holds. However this fragment spectrum poses an interesting question in terms of the structure elucidated for the related trimer. No mass losses indicating the loss of an ellagic acid are observed which then means that what was previously thought to be an ellagic acid, were 2 gallic acid units attached. As mentioned previously, the galloyl esters are the most probable positions for fragmentation of the oligomers. So far, the C-C bonds are stable against fragmentation and thus it can safely noted that the oligomers only consits of ellagic acid units from tetramers and the lower oligomers previously analysed only consist of long chains of gallic acid units.

This begs the question of wether the fragmentation between two linked galloyl units differs from an ellagic acid unit. Just looking at the structure one can observe that the galloyl ester bond is much more susceptible to fragmentation as compared to the C-C bond of the ellagic acid unit. The lower oligomers are then only composed of long chains of galloyl units and ellagitannins are only detected from tetramers.

The MALDI-CID fragmentations of the various oligomer peaks observed in the MALDI-TOF spectrum of the tara tannin provided unique information that allowed the specific microstructure of each oligomer to be determined unambiguously.

5.3.2.2. Turkey gall tannin

Application of the MALDI-CID technique for analysis of selected tara tannin oligomers presented a unique way to determine the microstrcuture. In order to compare the two polygallic structures observed for the tara and turkey gall tannins, selected ions from the latter were fragmented. The major difference between these two tannins is the basic unit to which the gallic acid units are added. In the case of the turkey gall tannin, the gallic acid units are still attached to the glucose monomer. As observed during analysis of the m/z 839 trimer from the tara tannin, the presence of the glucose monomer markedly alters the fragmentation pattern observed but also presents unique internal fragments that allow the location of the gallic acid units to be determined.

In order to investigate the structure of the turkey gall tannin in detail the m/z 811.6 ion belonging to a tetramer made up of 4 gallic acid units attached to glucose was analysed. The fragmentation spectrum is shown in **Figure 5.16.** As indicated in the figure the fragmentation pattern is relatively simple and resembles closely the type of fragmentation observed for the m/z 824 ion from the tara tannin (**Figure 5.9**). In the fragment spectrum of the m/z 811 ion there are no fragments resulting from the glucose monomer, unlike in the case of the m/z 839 ion from the tara tannin. This can be explained by considering the fragmentation behaviour of glucose oligomers. In this case only a single glucose monomer is present and thus higher energies would be required to form the 134 Da mass fragment previously shown. Also the glucose molecule due to its ring structure is more stable against fragmentation than the galloyl esters and therefore fragmentation preferentially occurs at these positions.

The fragmentation of the oligomer (m/z 811.6) continues by a sequential loss of three gallic acid units (**Figure 5.16**) to form the m/z 469.1 ion which represents 2 gallic acid units attached to the glucose monomer. A further loss of 125 Da from this ion leading to formation of the m/z 343.9 ion is observed. These fragments are specific to this oligomer and have not been observed for the other structures

present. For a similar type of fragmentation pattern in the case of the m/z 824 ion of the tara tannin the oligomer struture all the gallic acid units gave a mass of 152 Da with the chain-end giving 153 Da. Therefore the presence of the 125 Da mass loss in this spectrum suggests a slightly different structure which can be explained by the additional gallic acid unit being attached directly to the glucose monomer (**Figure 5.18**). The structure forming the m/z 811.6 ion can thus be assigned as 3 gallic acid units are linked to each other and another linked directly to the glucose monomer.

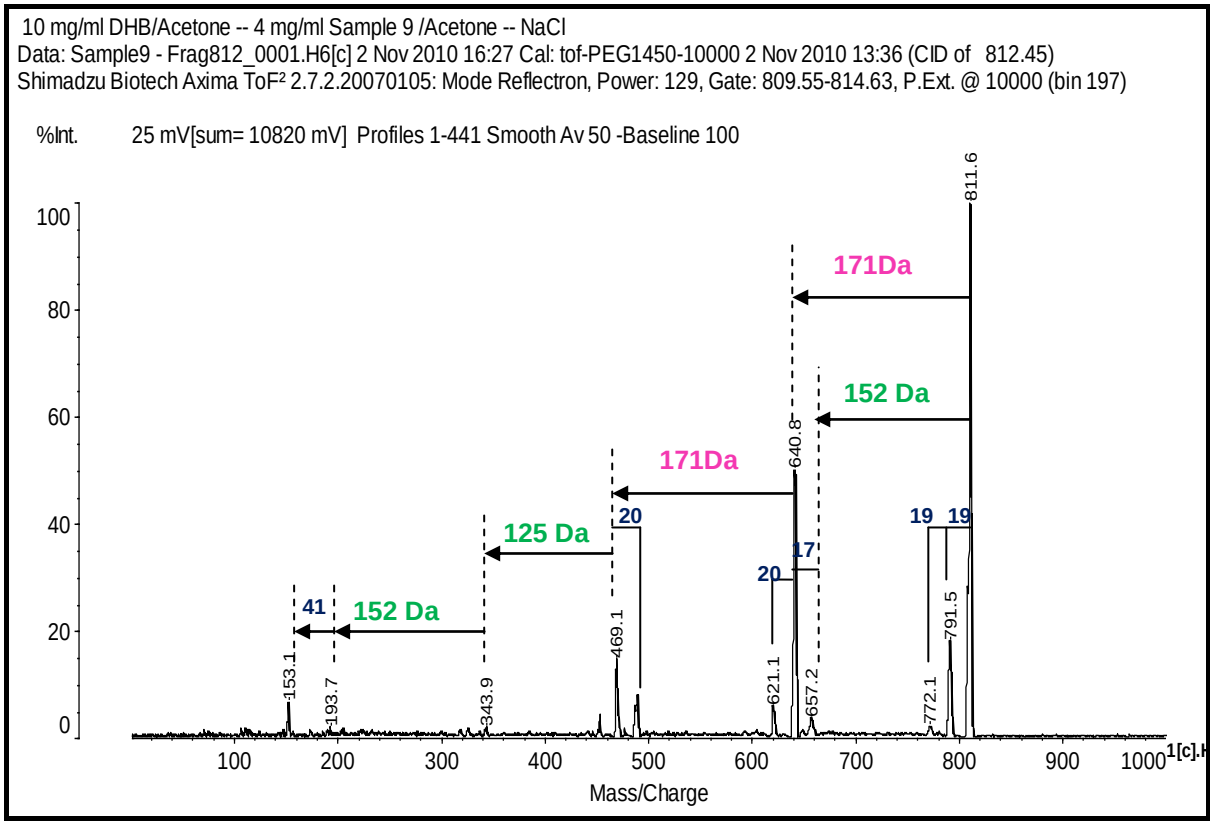


Figure 5.16: MALDI-TOF CID spectrum of the 811.6 Da trimer from solvent-extracted turkey gall tannin. The fragmentation pattern is illustrated by indicating the mass loss incurred from the precursor ion.

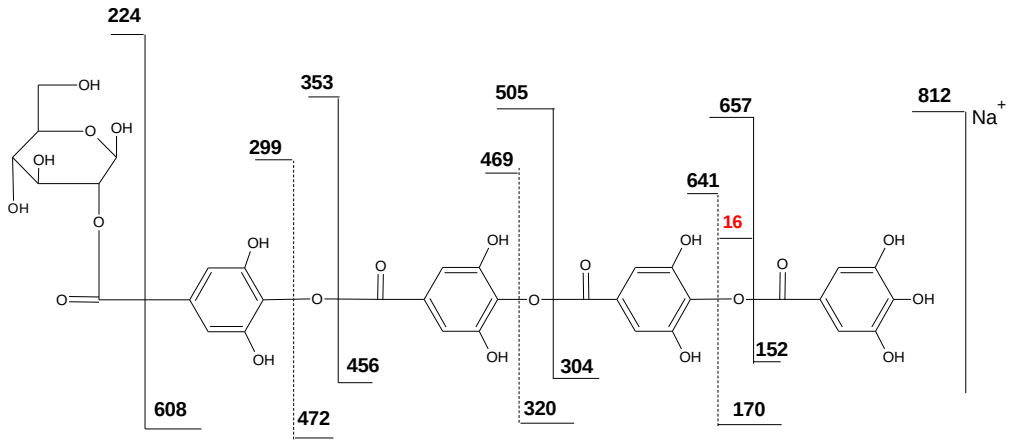


Figure 5.17: Initially proposed structure of the 811.6 Da trimer determined by MALDI-TOF of the turkey gall tannin.

Investigation of more complex oligomers was performed considering that the fragmentation behaviour of various oligomer structures was determined. The m/z 964.2 ion observed in the MALDI-TOF spectrum of the turkey gall tannin has been assigned to pentagalloyl glucose (PGG) (**Figure 5.20**). However, from the analysis that has been carried out thus far it was shown it is not possible from the MALDI-TOF spectrum to determine the location of the gallic acid units. In order to investigate the structure, the m/z 964.2 ion was fragmented and the resultant MALDI-CID spectrum is shown in **Figure 5.19**. As indicated in this figure there is a sequential loss of 170 Da mass, this is the same type of behaviour observed in the fragmentation of the m/z 811.6. The same fragment m/z 469 ion (observed as m/z 471) is present and was assigned to the glucose monomer attached to 2 gallic acid units. A mass of 160 Da is then lost from this fragment ion in order to form the m/z 311.8 peak. This peak is a sodium adduct of the m/z 288 fragment which was not detected in the case of the m/z 839 ion from the tara tannin. This means that internal fragments from the glucose ring are formed, and this structure is consistent with oligomer chains whereby the galloyl units are attached directly to the glucose. The 152 and 170 Da mass losses are result of the gallic acid units being stripped from the glucose core. Therefore the structure that was postulated is confirmed as a PGG monomer.

Further investigations were performed on a pentamer related to the PGG monomer by 17 Da. As shown for the tara tannin, a small difference in masses between related oligomers can display very different structures. This point is further illustrated by analysis of the m/z 947.2 ion, the MALDI-CID spectrum is shown in **Figure 5.21**. The fragment ions are exactly the same ones observed for the m/z 811.2 ion, this oligomer just has an additional gallic acid unit. The additional group is added to one of the gallic acid units on the structure since its mass is 152 Da.

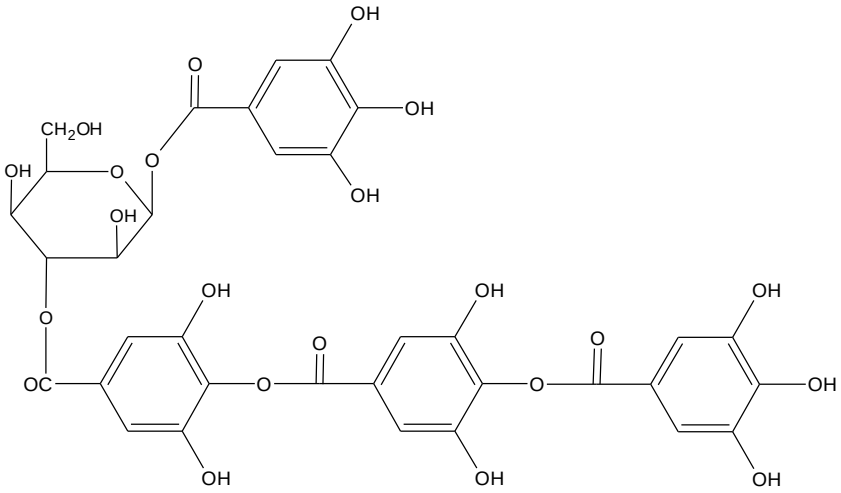


Figure 5.18: New structure proposed for the oligomer appearing at 811.6 Da following MALDI-CID analysis which allowed oligomer sequence determination.

In order to support the conjectures that were made during structure elucidation further investigation of a different oligomer was analysed. The m/z 846.8 ion was analysed and the fragmentation spectrum is presented in **Figure 5.22**. Hardly any fragments are forming from this oligomer, although from the MALDI-TOF spectrum its structure seemed related to the m/z 811.6 oligomer. Clearly the two structures are not the same, although they are both be trimers. The precursor ion for this oligomer is not detected in the MALDI-CID but rather a peak is observed at m/z 842. In the case where the gallic acid units are attached to each other or the glucose some fragmentation is observed Therefore from this perculiar result it can be assumed that this the m/z 846.8 ion is formed by a complex oxidation product that would

require higher energies than those available in the CID in order to form fragments. Molecules such as castalin/vescalin have been detected in some tannins and this assumption would need further investigation by analysing such compounds by MALDI-CID and observing their behaviour.

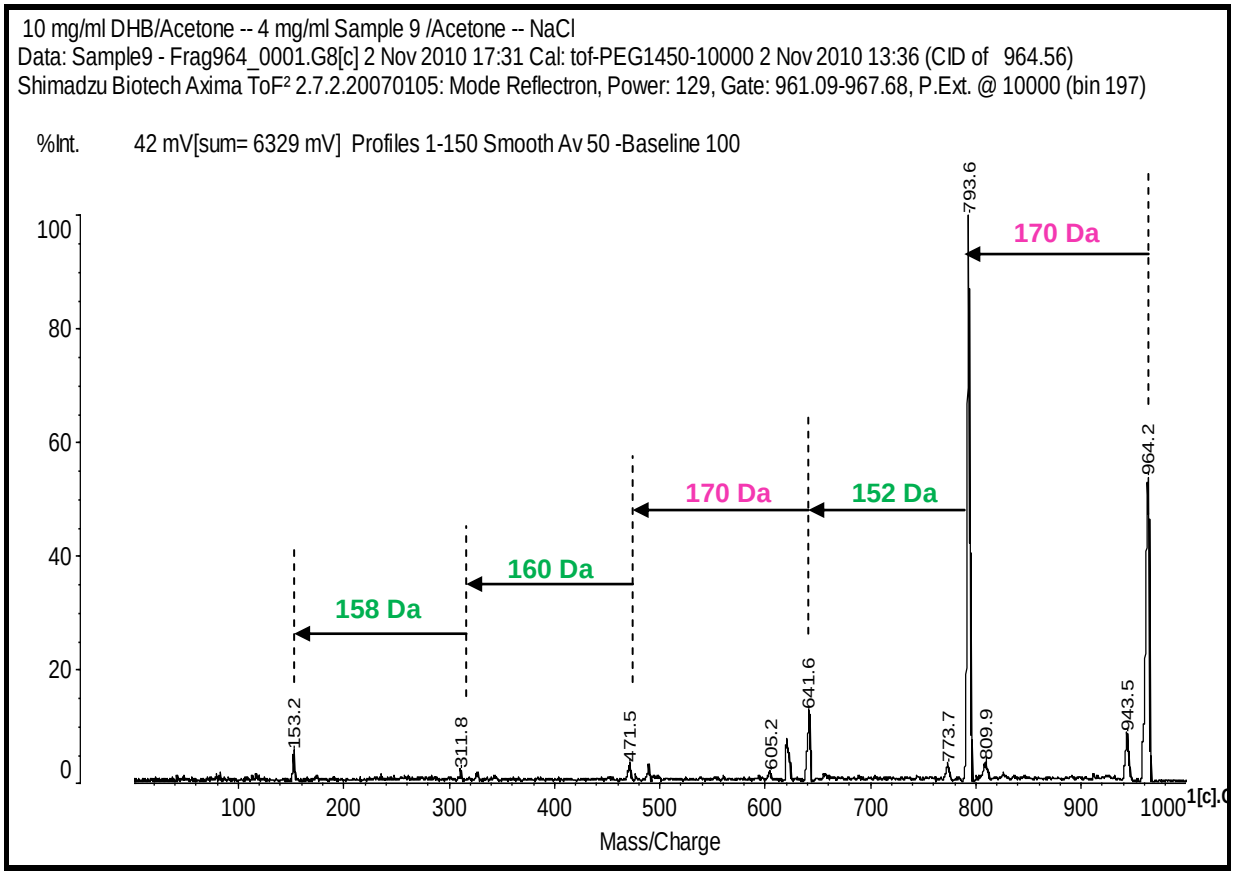


Figure 5.19: MALDI-TOF CID spectrum of 964.2 Da peak representing a PGG from solvent-extracted turkey gall tannin.

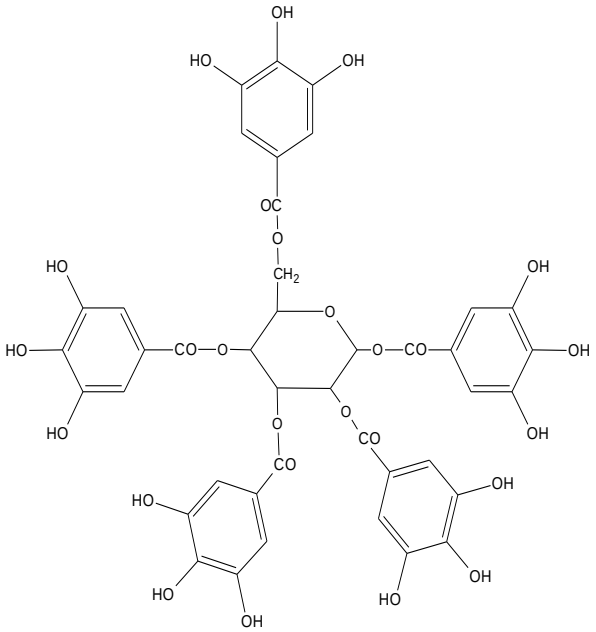


Figure 5.20: Structure of pentagalloyl glucose (m/z 963)

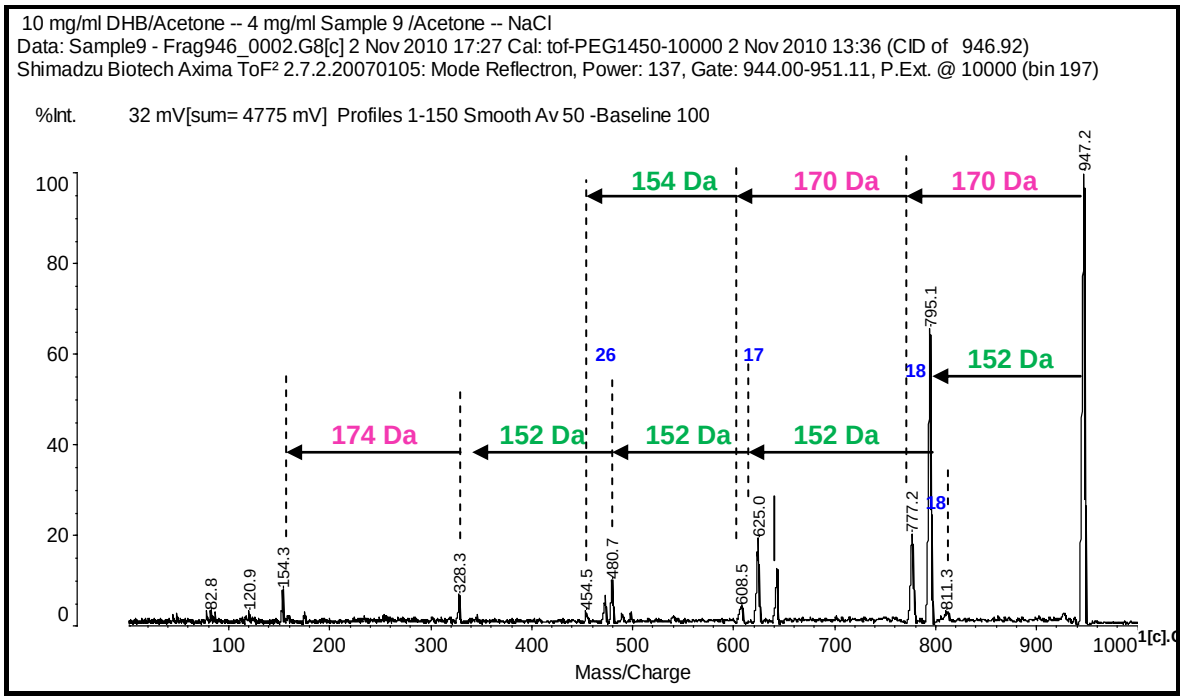


Figure 5.21: MALDI-TOF CID spectrum of 947.2 Da peak representing a pentamer from solvent-extracted turkey gall tannin.

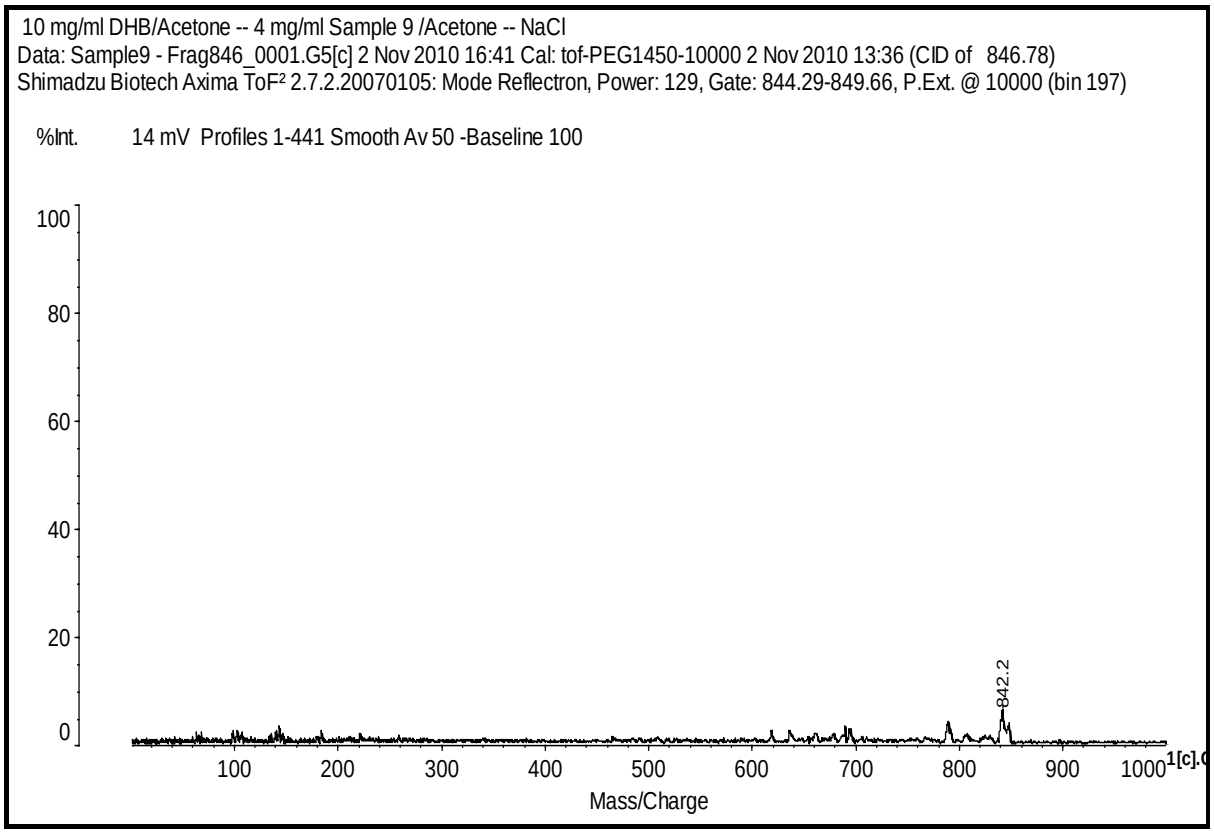


Figure 5.22: MALDI-TOF CID spectrum of 842.2 Da peak representing a tetramer from solvent-extracted turkey gall tannin.

5.3.2.3. Chestnut tannin

The chestnut tannin is composed of ellagitannins and polygallic oligomers. The main constituents in this solvent-extracted tannin are galloyl unit oligomers stripped from their glucose core. The applicability of the MALDI-CID has shown to be able to distinguish between the various oligomer structures and thus it will be applied here in order to note the sequence distribution of the oligomers in this extract. The fragmentation spectrum for the oligomer peak at 672.3 Da is shown in **Figure 5.23**.

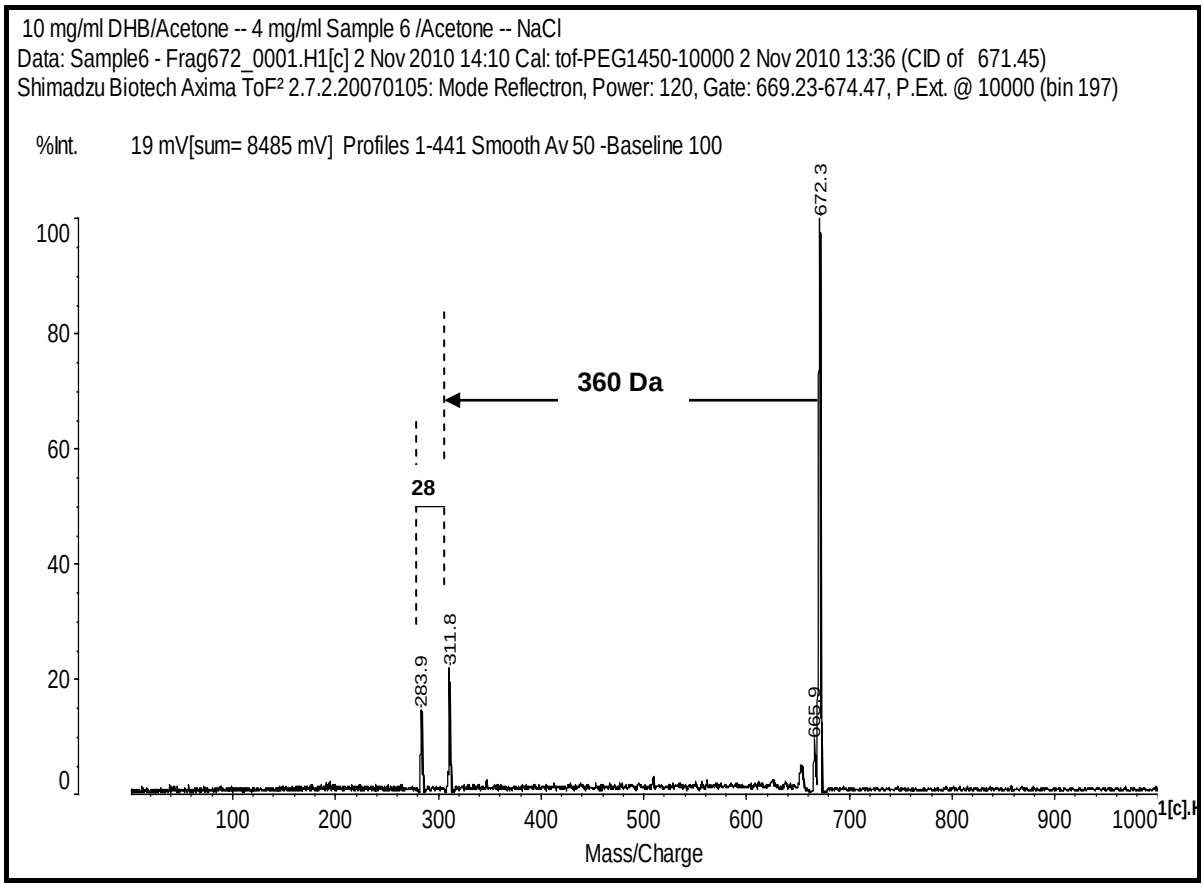


Figure 5.23: MALDI TOF CID spectrum of the 672.3 peak representing a dimer (digalloyl glucose attached to dehydrated glucose) from solvent-chestnut tannin.

There are only two fragments formed when the m/z 671 oligomer ion is analysed, the postulated structure based on MALDI-TOF analysis is represented in **Figure 5.24**. The fragmentation patterns have been discussed at length for the two polygallic tannins. The fragment spectrum for the oligomer in the chestnut tannin extract does not resemble any of the previously discussed spectra. However, the m/z 311 ion was also detected in the case of the MALDI-CID analysis of the PGG monomer from the tara tannin. The fragment mass that is lost by the precursor ion is not consistent with gallic acid units linked via a galloyl ester bond. The most likely structure represented by the 360 Da fragment is a gallic acid unit that is attached the CH_2-O- of the glucose monomer (**Figure 5.25**). This structure is clearly related to the PGG monomer with the other gallic acid unit remaining attached to the glucose structure and the location on the glucose monomer cannot be determined.

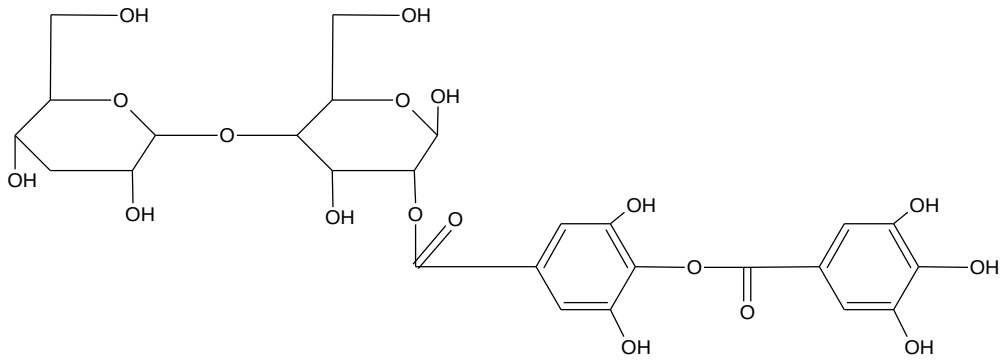


Figure 5.24: Proposed structure for the peak at 671 Da representing a dimer (digalloyl glucose attached to dehydrated glucose) present in solvent-extracted chestnut tannin. The structure was determined according to bulk MALDI-TOF analysis of the tannin.

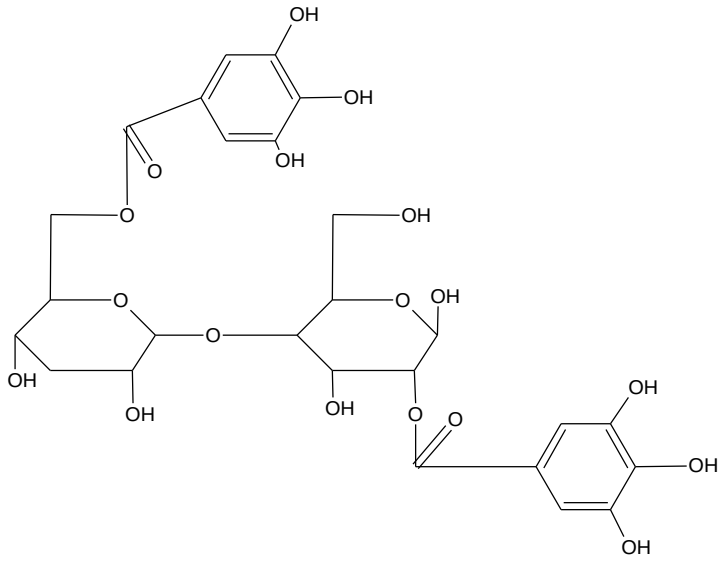


Figure 5.25: Structure of dimer with 672.3 Da mass determined by MALDI-TOF CID analysis.

MALDI-TOF CID analysis provides unique structural information and gives an insight into the oligomer microstructure in a relatively simple way. Analysis of the hydrolysable tannins by this method not only provides information on the location of the gallic acid unit but in the cases where its applicable the oligomer monomer sequence was also determined. Therefore presented in this section is a new mass

spectrometric technique that is able to determine the distribution and location of the gallic acid units in addition the linking positions in the oligomers has been developed.

5.4. Determination of chemical composition by ^{13}C NMR

The chestnut and tara tannins were analysed by ^{13}C NMR. The samples were selected due to their different chemical structures. Determination of the chemical structure by NMR is a well developed method in the case of hydrolysable tannins [18-20]. For specific structural determinations information obtained from both ^{13}C NMR and ^{1}H spectra is utilised. Normally, correlation spectra are used in order to confirm structural assignment [20]. In order to determine structure specific molecules are isolated, methylated and then analysed [18].

In this study the bulk samples were analysed only by ^{13}C NMR and as can be seen the spectra show resonances in the same regions. The chestnut tannin sample is more complex compared to the gallotannins as indicated by the MALDI-TOF analysis carried out. The peak assignments were completed by comparing the chemical shifts to those obtained in the analysis of pure trimers and tetramers obtained from literature [18]. In the spectrum for the chestnut tannin the broad multiplet of peaks appearing between 60 and 80 ppm belongs to the carbons on the glucose moiety, that is C1, C2, C5, C6 and C4 (**Figure 10**). The presence of the four well-resolved peaks between 85 and 103 ppm indicates the presence of four glucose units present in the extract. The presence of these peaks thus also confirms the data obtained from MALDI-TOF that the chestnut extract is composed of long chains of glucose units attached to galloyl units. The presence of chemical shifts belonging to the glucose moieties even after solvent extraction which removes the carbohydrate fraction of the tannin indicates that the carbohydrate in this case forms part of the polymer chain. The other peaks for the glucose resonances are located between 63 and 76 ppm, these indicate the C6, C5 and C6 carbons. Depending on the substitution of the adjacent carbons the chemical shifts may vary slightly.

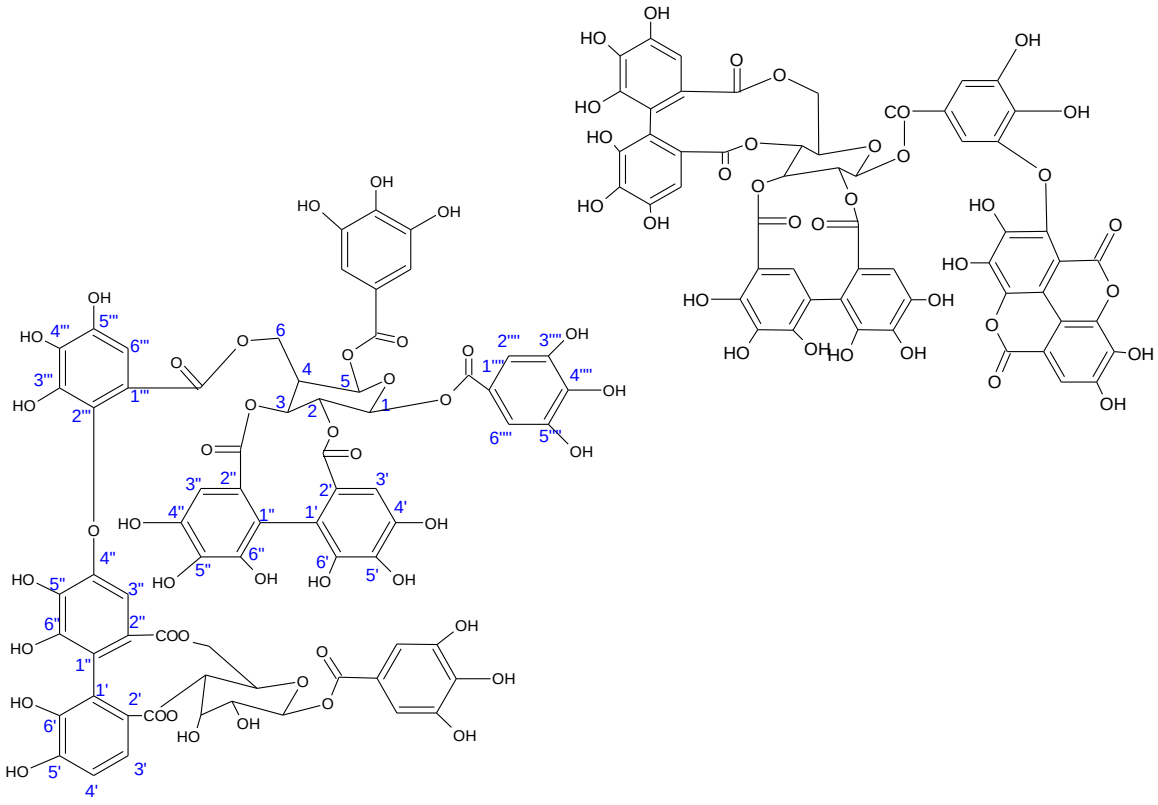


Figure 5.26: Hydrolysable tannin structures used as a models for ^{13}C NMR peak assignments. (a) Nobotanin R [20] (b) davuriciin M_1 with valoneoyl group isolated from the root of Rosa davurica. [19].

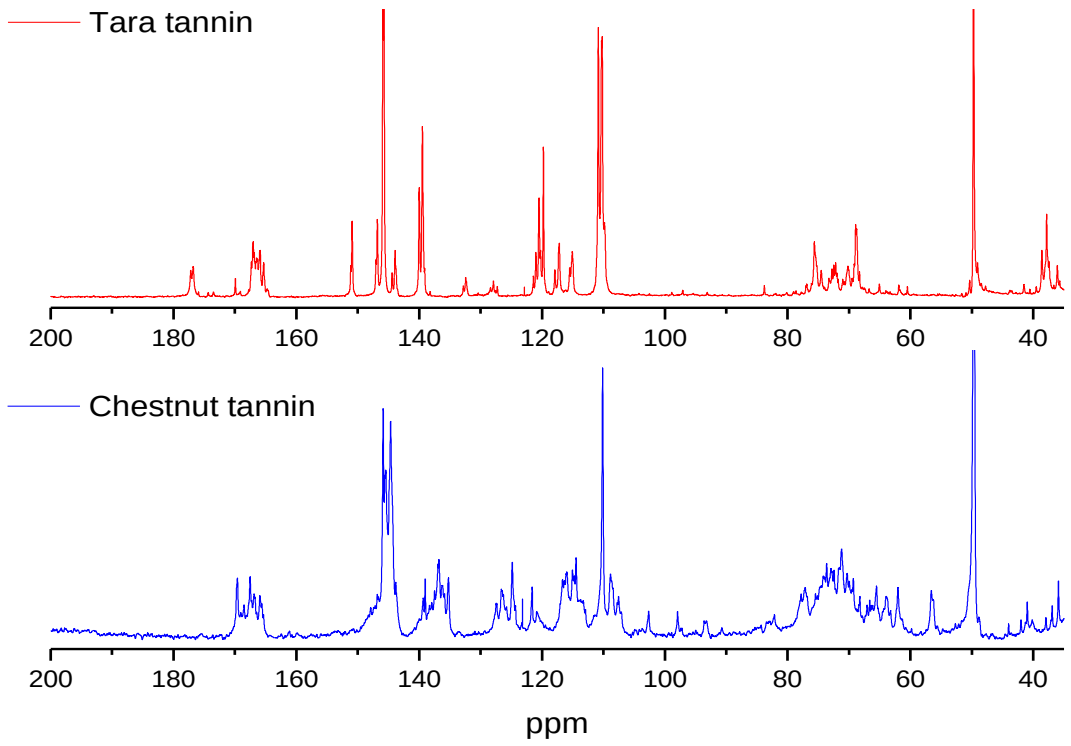


Figure 5.27: ^{13}C NMR spectra of tara tannin and chestnut tannin. Chemical shift range shown from 200-35 ppm, the aliphatic region is not shown since it overlaps with the solvent (d_6-acetone) signal.

Above 100 ppm are located the signals of the phenolic moieties attached to the glucose cores. The ^{13}C NMR spectrum of the chestnut tannin also reveals that there are some HHDP (ellagic acid) units on the glucose moieties; this is indicted by the presence of peaks at 107 and 108 ppm (C3', C3") [20]. These peaks are not observed in the spectrum of the tara tannin. At 110 ppm a sharp peak is observed in the spectrum of both tannins, two peaks are observed 110.1 and 110.8 ppm, these peaks belong to the gallic acid units (C2', C4', and C6') [18,20]. Another type of structure that is present and whose peaks are observed in the case of the chestnut tannin is the valoneoyl structure [20] (**Figure 5.26**). The presence of this group is indicated by the signals appearing at 102 ppm belongs to the C3'' on this structure and the C6''' is at 109 ppm. The signals for the vaneoyl C3' and C1''' and HHDP C1', C1''

appear around 114 ppm and this appears as a broad band in the spectrum, indicative the overlap that is occurring. For the tara tannin this region is less complex, with two distinct signals that can be believed to belong to the C2 of the gallic acid unit. The peaks in the region between 120 and 130 ppm belong to the HHDP and vaneoyl group C2 carbons. In the tara tannin these peaks can be assigned to some of the C1 carbons on the gallic acid moeity. The C3 and C5 of this group are located at 140 ppm as two distinct signals and the C4 is at 146 ppm. The spectrum of the chestnut tannin however is not as simple, these peaks belonging to the galloyl moiety overlap with the C3''', C4', C4'', C4''' and C5''' of the vaneoyl group. The C6', C6'' and C4', C4'' of the HHDP structure also appear in this region, further complicating the spectrum. The C7 of all the phenolic groups appears in the region of 168 ppm.

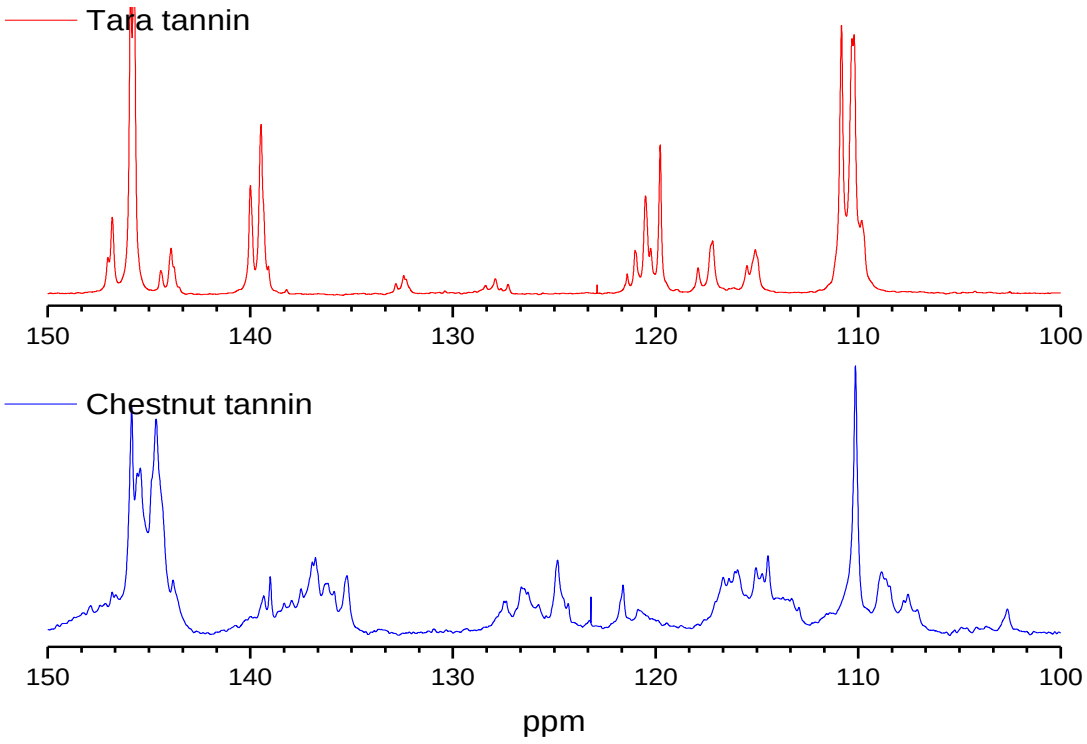


Figure 5.28: ^{13}C NMR spectra of tara tannin and chestnut tannin. The chemical shift range is shown from 150-100 ppm.

As observed in the structure the various phenolic moieties may be located at any of the free OH groups on the glucose structure. Further, there exists long chains of these glucose structures in the case of the chestnut tannin, this means that there are even more possibilities in terms of where the phenolic groups can attach to the oligomeric structure. In the case of the tara tannin, the polymerisation pattern is simple enough which means the oligomers present in the extract are relatively similar and this leads to a less complex ^{13}C NMR spectrum. The complexity in the structure of the chestnut tannin is added by the various oxidation products such as vaneouyl and HHDP which have signals in the same region but small shifts occur due to the different chemical environments surrounding them.

In order to obtain even more information on the location of the different structures along the glucose chains, especially in the case of the chestnut tannin, coupling experiments in NMR are required. Although the ^{13}C NMR analysis gave an idea of the structures that are present, ^{1}H NMR and COSY experiments are required in order to combine this information [20]. The crude samples are analysed in this case and thus there are various structures overlapping and this makes assignment difficult. Analysis of high molar mass tannins give broad peaks in ^{13}C NMR and this make analysing them by this method challenging. However, the purpose here was to show that ^{13}C NMR can be used as a tool in the analysis of crude commercial tannin extracts in a comparative manner. A qualitative survey of the presence and/or absence of certain characteristic chemical shifts can lead to positive identification of the components of that particular extract. In combination with MALDI-TOF the tara and chestnut tannin extracts main constituents were confirmed. Using ^{13}C NMR revealed the presence of a vaneoyl group whose presence was not detected in MALDI-TOF, this could be due to the fact that this structure would show a mass difference of an ellagic acid unit + a single galloyl unit. MS techniques are unable to distinguish between such structures and thus ^{13}C NMR and MALDI-TOF MS were shown to be complementary techniques in this case.

5.5. Molar mass distribution determination by SEC

The analysis of hydrolysable tannins by SEC is not widespread. Yanagida et al. showed an optimised SEC method that separated tannic acid which is a mixture of low molar mass gallotannins and some ellagitannins [11]. The analysis method proved successful in separating both condensed and hydrolysable tannins. In view of this fact, the method that was developed for the condensed tannins (adapted from Kennedy et al. [16]) was used in application to the chestnut and tara tannin [16].

Table 5.5: Description of tannin samples and the average molar mass† determined by SEC making use of a PMMA calibration†.

Sample	Description	M_n	M_w	PDI	DP
6	Solvent-extracted chestnut	11200	12898	1.2	11
7	Solvent-extracted tara	6000	8700	1.5	12

†*the molar masses are relative to narrow PMMA standards*

The tannin samples were analysed in the same mobile phase as the condensed tannins in order to determine their molar masses. For each of the tannin extract the degree of polymerisation, Mn and Mw were determined, these results are shown in **Table 5.5**. The tara tannin (sample 7) was used to determine whether the SEC separation would be suitable for hydrolysable tannins. The resultant chromatogram is shown in **Figure 5.29** and the separation of oligomers is observed in the lower molar mass region, and then the major peak has a shoulder. The presence of lower molar mass components and broadness of the peak at higher molar mass indicate the chemical heterogeneity of this extract. This is known fact for tannin extracts, they contain a variety of structures that have a wide molar mass distribution. The end of separation reaches only up 10 000 Da which means this the molar mass

maximum for this tannin. As mentioned in previous chapter care must be taken to assign masses as seen in the chromatogram since a Universal calibration was used.

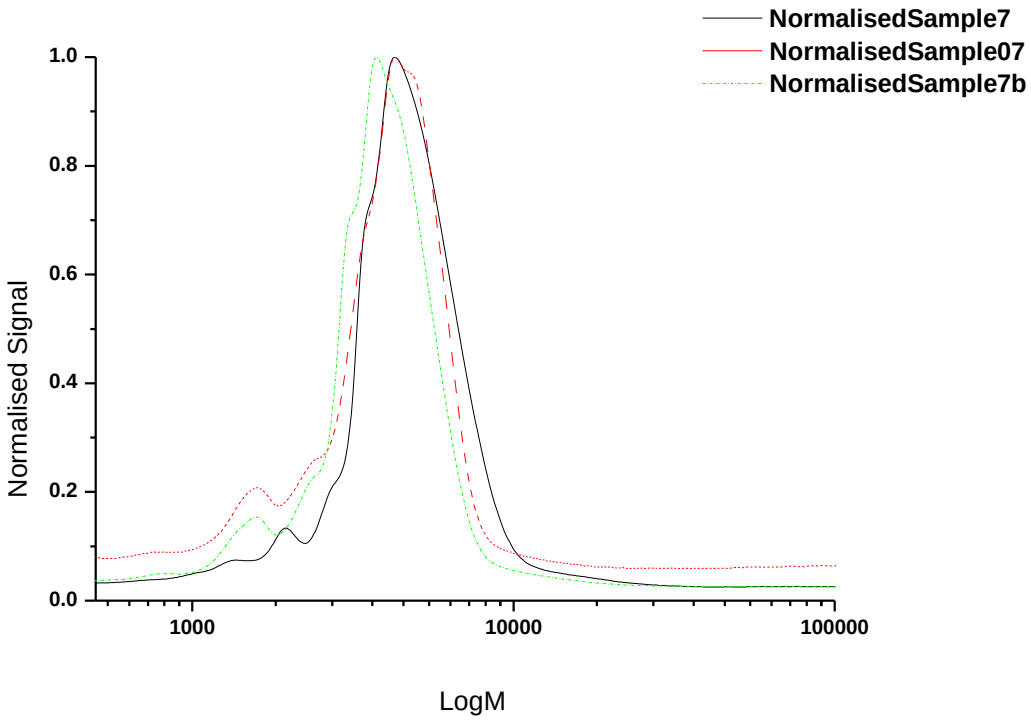


Figure 5.29: Signal response vs. LogM plots of (10 mg/mL) tara tannin (sample 7) with different injection volumes. Black: 50 µL injection volume. Red: 20 µL injection volume. Green: 20 mg/mL sample, injection volume: 20 µL.

The chestnut ellagitannin was also analysed by SEC. The chromatogram is shown in **Figure 5.30**, the 2 peaks indicate different injection volumes, and clearly separation is not affected by the concentration. When compared to the gallotannin the chromatogram is very different. Only a single peak at approximately 2 000 Da is observed for the lower molar mass. This means that this tannin is mainly composed of oligomers at 6000 Da. The shoulder on the higher molar mass region and the tailing in this shows the presence of higher molar masses. The tailing may indicate the presence of some interactions with the stationary phase. In order to investigate this phenomenon, this fraction of the oligomeric

separation can be analysed by MS techniques such as MALDI-TOF in order to observe the components. However, this is challenging since LiCl is added in significant amounts in order to aid the separation, in the concentration step the sample is saturated by salt which makes further analysis impossible. In order to circumvent this problem a different mobile phase composition would be required in order to analyse the molecules. An overlay of both the chestnut and tara tannin are overlaid in **Figure 5.31**.

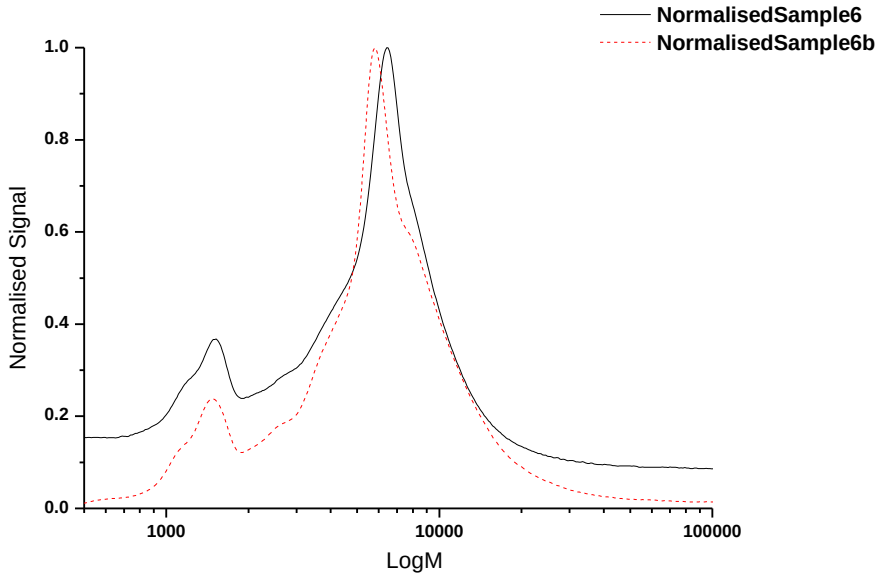


Figure 5.30: Signal response vs. LogM plots obtained by SEC analysis (calculated using the PMMA calibration) of (10 mg/mL) chestnut tannin (sample 6) with different injection volumes. Red: 50 µL injection volume. Black: 20 µL injection volume. Stationary phase: 2 x PLgel mixed- D columns (300 × 7.5 mm i.d., 5 µm) connected in series and protected by a guard column with the same material. Mobile phase: DMAc with 0.03 % (w/v) LiCl and 0.05 % (w/v) BHT. Flow rate: 1 mL/min.

Both the chestnut and tara tannin extracts seem to be relatively low molar mass extracts compared to their proanthocyanidin counterparts.

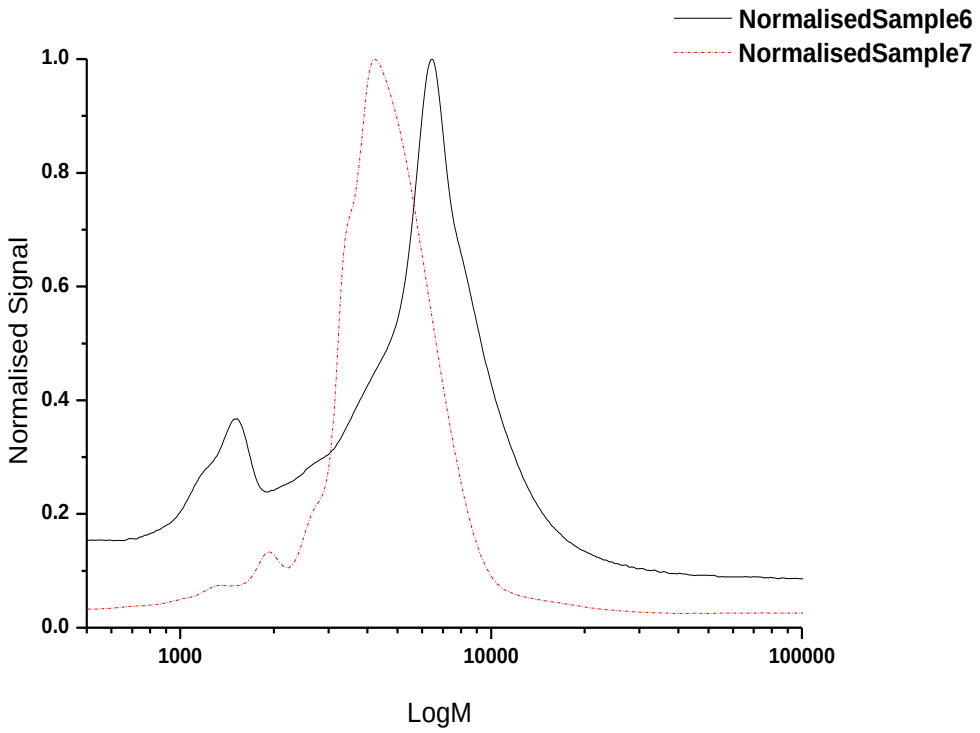


Figure 5.31: Overlay of signal response vs. LogM plots (calculated using the PMMA calibration) obtained by SEC analysis chestnut and tara tannins, at the same concentration. Stationary phase: 2 x PLgel mixed- D columns (300 × 7.5 mm i.d., 5 µm) connected in series and protected by a guard column with the same material. Mobile phase: DMAc with 0.03 % (w/v) LiCl and 0.05 % (w/v) BHT. Flow rate: 1 mL/min.

5.6. Separation of oligomers by normal phase chromatography

The separation of hydrolysable tannins according to degree of polymerisation is not as well studied as in the case of proanthocyanidins. The separation mode of choice for hydrolysable tannins is reverse phase liquid chromatography, which was shown to separate the molecules according to chemical composition in the same way as observed for the oligomeric proanthocyanidins [12]. The information on the

separation of oligomeric hydrolysable tannins by normal phase chromatography is limited and although this is the case some attempts have been made to analyse these extracts [1,7,19,21]. The method used by Hagerman to separate hydrolysable tannin extracts from commercial tannic acid and fireweed flower tannin showed good separation of the low molar mass tannic acid, however, for the fireweed which seemed to contain higher molar masses the chromatographic system was unable to separate the higher molar mass fraction [1]. The mobile phase used was composed of THF, methanol and hexane and was able to separate the molecules to some degree and thus warranted further investigations.

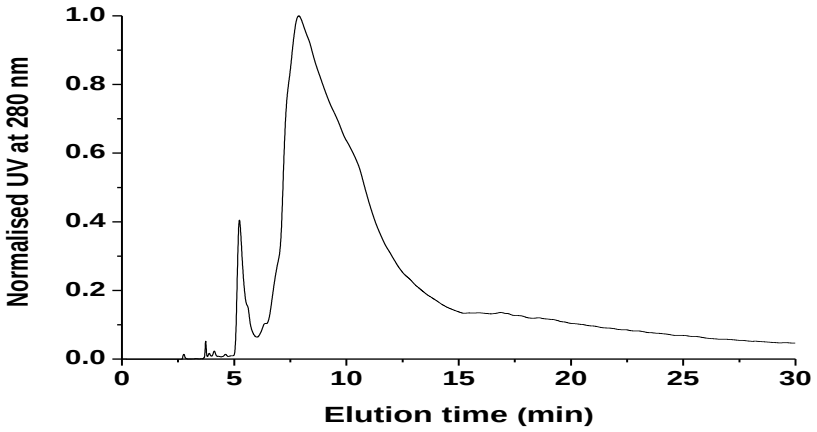


Figure 5.32: Isocratic NP-LC separation of solvent-extracted tara tannin (sample 7), mobile phase: 58 % n-hexane and 42 % solvent B (3:1 methanol/THF and 0.25 % acetic acid), stationary phase: Nucleosil unmodified silica (150 x 4.6 mm, i.d. 5 μm), flow rate: 1 mL/min, temp: 30°C. Detector: UV at 280 nm.

In order to develop a chromatographic method that would be able to separate the higher molar mass tannins considered in this study the mobile phase composition was kept at 58% (v/v) hexane and 42 % solvent B at a flow rate of 1 mL/min [1]. Solvent B is composed of 3:1 methanol/THF and 0.25% acetic acid. The chromatogram showed in **Figure 5.32** was obtained for the tara tannin, by isocratic elution with the afore mentioned mobile phase composition. The separation obtained was not good as seen in

the chromatogram (**Figure 5.32**). In order to improve separation a linear gradient was first attempted, however this improved the separation only slightly **Figure 5.33a**. The separation at this point as seen in **Figure 5.33a** resembled the chromatograms observed for the fireweed flower tannin that is, at higher elution volumes there was no separation. A closer look at the chromatogram reveals some form of separation, although not resolved this indicated that the separation may be improved further.

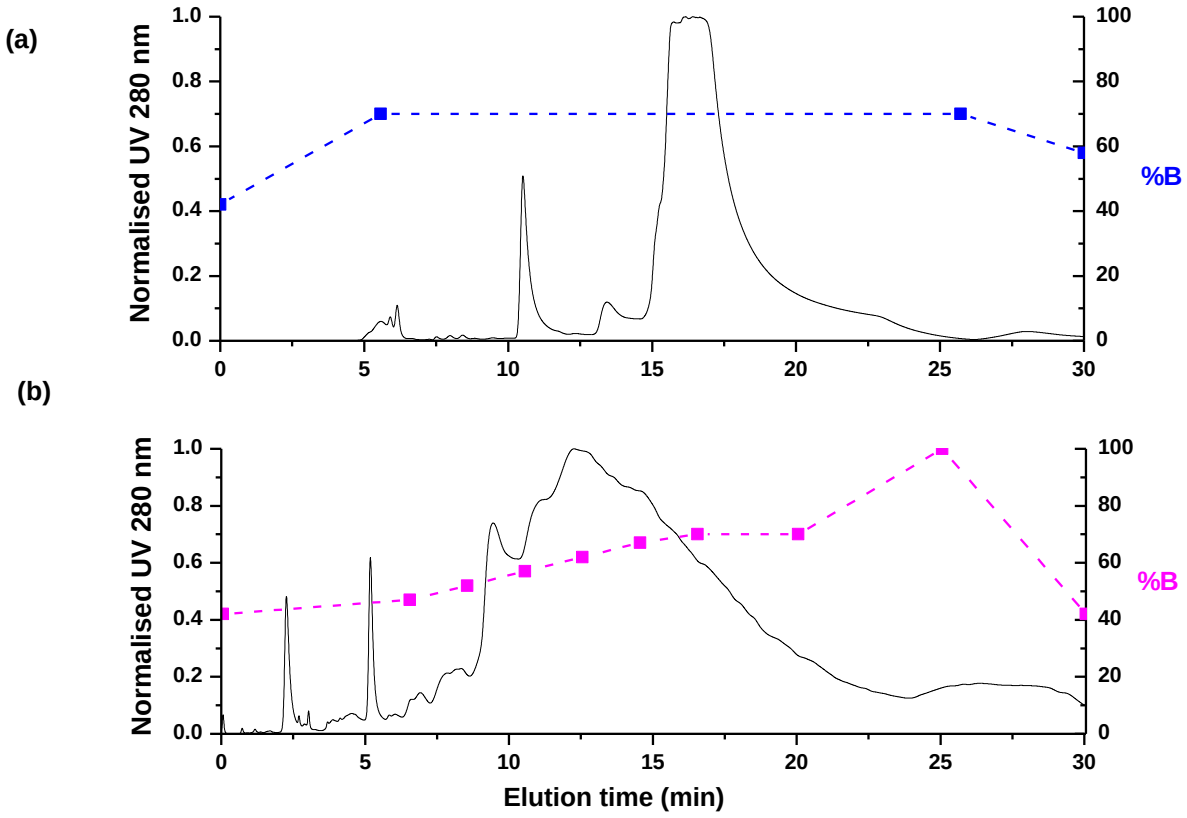


Figure 5.33: Gradient NP-LC separation of solvent-extracted tara tannin (sample 7), initial mobile phase composition: 42 % n-hexane and 58 % solvent B (3:1 methanol/THF and 0.25 % acetic acid), stationary phase: Nucleosil unmodified silica (150 x 4.6 mm, i.d. 5 μm), flow rate: 1 mL/min, temp: 30°C. Detector: UV at 280 nm. The dotted lines indicate mobile phase composition in the gradient.

In order to improve separation, a form of step gradient increasing the amount of non-solvent slower was introduced and the optimised separation is shown in **Figure 5.33b**. The chromatograms obtained with the respective gradients were plotted with the solvent composition of solvent B indicated in the right hand side axis **Figure 5.33.** The optimised separation of tara tannin showed good repeatability (**Figure 5.34**). The sample was injected in different concentrations in order to test its validity. The samples with a lower concentration show a lower absorbance as expected, this did not alter the elution of the molecules though. The gradient that was used for further analysis is the one indicated in **Figure 5.33b**. It was also of interest to check the effect of concentration on separation in this case since further coupling with spectroscopic techniques may require fractionation.

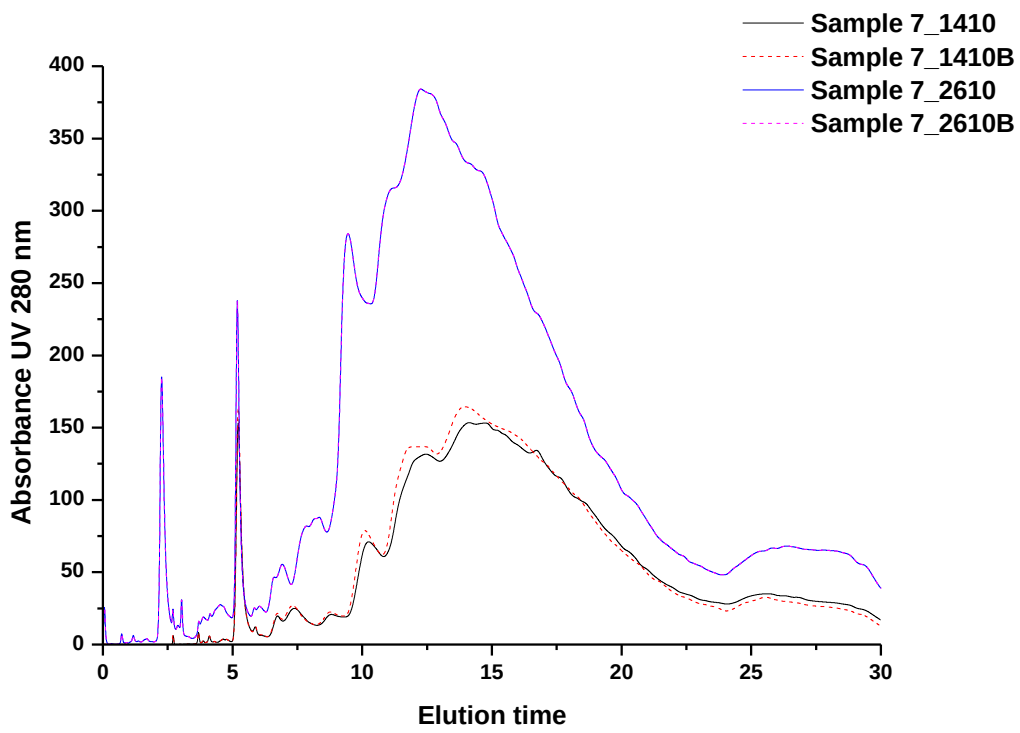


Figure 5.34: Repeatability study of Gradient NP-LC separations of solvent-extracted tara tannin (sample 7) at different concentrations and injection times using the optimised gradient indicated in Figure 5.33b. The numbers represented in the legend indicate the injection times by date and month.

Assuming that the separations that occurred were according to size, it can be clearly seen that of the two tannins the tara tannin contains the higher molar mass constituents. The intensity of the peak at earlier elution time in the case of the turkey gall tannin indicates that the highest proportion of this extract is low molar mass compounds although some increased molar mass constituents are also present. The point of developing a chromatographic technique was such that hydrolysable tannins of different chemical composition may be analysed and compared. In order to achieve this goal, the turkey gall tannin was analysed using the optimised gradient and the chromatogram showed in **Figure 5.35** was obtained. As indicated by this chromatogram further adjustments are required in order to accommodate the difference in structure of these two tannins. The tara and turkey gall tannins are both gallotannins, therefore need to change the separation indicates the sensitivity of NP-LC separations to structural differences.

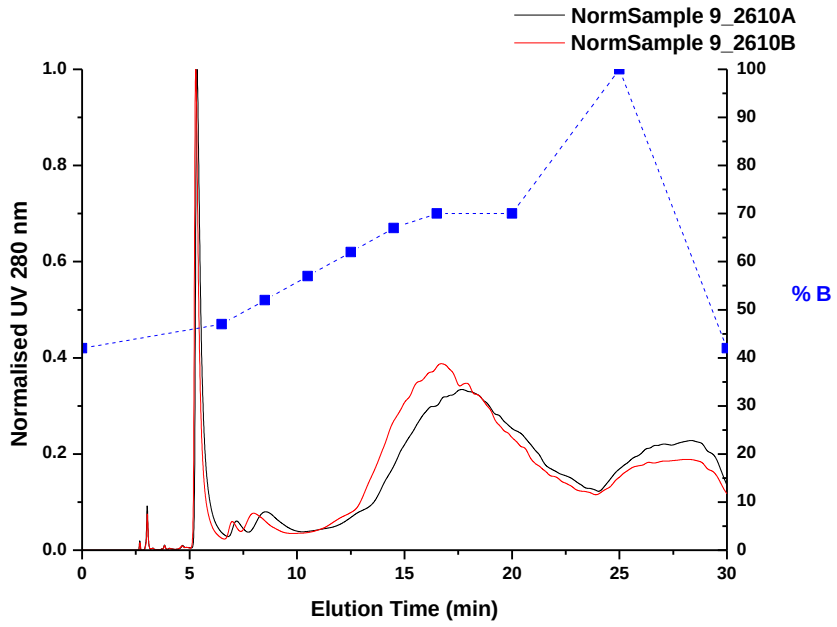


Figure 5.35: Gradient NP-LC separations of solvent-extracted turkey gall tannin (sample 9), initial mobile phase composition: 42 % n-hexane and 58 % solvent B (3:1 methanol/THF and 0.25 % acetic acid), stationary phase: Nucleosil unmodified silica (150 x 4.6 mm, i.d. 5 µm), flow rate: 1 mL/min, temp: 30°C. Detector: UV at 280 nm. The dotted lines indicate mobile phase composition in the gradient.

Unfortunately the chestnut tannin could not be analysed by this mobile phase composition due to its insolubility. This further illustrated the disadvantage of using interaction chromatography for size separations. The NP-LC method investigated in this study shows good separations of tannins that are made up of gallotannins, however another method that can analyse both gallo- and ellagitannins is required.

5.7. References

[1] A.E. Hagerman, C.T. Robbins, Y. Weerasuriya, T.C. Wilson, C. McArthur, J. Range Manage. 45 (1992) 57.

[2] A.E. Hagerman, in A.E. Hagerman (Editor), Tannin Handbook, http://www.users.muohio.edu/hagermae/tannin.pdf, Oxford, 1998.

[3] A. Pizzi, H. Pasch, K. Rode, S. Giovando, J. Appl. Polym. Sci. 113 (2009) 3847.

[4] H. Pasch, A. Pizzi, J. Appl. Polym. Sci. 85 (2002) 429.

[5] A. Pizzi, in M. Belgacem, A. Gandini (Editors), Monomers, Polymers and Composites from Renewable Resources, Elsevier, 2008, p. 179.

[6] E. Haslam, Practical polyphenolics from structure to molecular recognition and physiological action, Cambridge University Press, New York, 1998.

[7] I. Mueller-Harvey, Anim. Nutr. Feed Technol. 91 (2001) 3.

[8] J.-P. Salminen , V. Ossipov, J. Loponen, E. Haukioja, K. Pihlaja, J. Chromatogr. A 864 (1999) 283.

[9] A. Romani, F. Ieri, B. Turchetti, N. Mulinacci, F.F. Vincieri, P. Buzzini, J. Pharm. Biomed. Anal. 41 (2006) 415.

[10] B. Zywicki, T. Reemtsma, M. Jekel, J. Chromatogr. A 970 (2002) 191.

[11] A. Yanagida, T. Shoji, T. Kanda, Biosci. Biotechnol., Biochem. 66 (2002) 1972.

[12] A. Yanagida, T. Kanda, T. Shoji, M. Ohnishi-Kameyama, T. Nagata, J. Chromatogr. A 855 (1999) 181.

[13] P. Mämmelä, H. Savolainen, L. Lindroos, J. Kangas, T. Vartiainen, J. Chromatogr. A 891 (2000) 75.

[14] C. Viriot, A. Scalbert, C.L.M. Herve du Penhoat, C. Rolando, M. Moutounet, J. Chromatogr. A 662 (1994) 77.

[15] N. Vivas, M.-F. Nonier, N.V. de Gaulejac, I.P. de Boissel, C. R Chimie 7 (2004) 945.

[16] J.A. Kennedy, A.W. Taylor, J. Chromatogr. A 995 (2003) 99.

[17] E. Spina, L. Sturiale, D. Romeo, G. Impallomeni, D. Garozzo, D. Waidelich, M. Glueckmann, Rapid Communications in Mass Spectrometry 18 (2004) 392.

[18] M.S.A. Marzouk, Phytochem. Anal. 19 (2008) 541.

[19] T. Yoshida, Z.-X. Jin, T. Okuda, Phytochemistry 30 (1991) 2747.

[20] J.H. Isaza, H. Ito, T. Yoshida, Phytochemistry 65 (2004) 359.

[21] T. Okuda, T. Yoshida, T. Hatano, J. Nat. Prod. 52 (1989) 1.

Chapter 6

Summary, Conclusions and Future work

In recent years the drive towards the use of 'green' materials has led research in the direction of making full use of naturally available compounds. Tannin molecules have been discovered and extracted from various plant species. Their main attraction is the health benefits they provided, however, these molecules are also important commercially. The tannin molecules differ in chemical composition and molar mass depending on which plant or extraction mode was used in obtaining them [1,2]. Although a great amount of research has been carried out, full characterisation of high molar mass tannin extracts has not been possible. However, in order to fully understand the structure-property relationships of these molecules, optimisation of analytical methods able to elucidate specific structures is crucial.

The aim of this thesis was to develop analytical methods and techniques to characterise commercial tannin extracts. Selective commercial condensed and hydrolysable tannins obtained from various plant sources were analysed. The main focus was on the combination of information obtained from spectroscopic techniques which provide chemical composition information with chromatographic separations aimed at separating the tannin molecules according to the degree of polymerisation.

The tannin samples of interest were divided into two main subgroups; (i) condensed tannins (proanthocyanidins) and (ii) hydrolysable tannins. The main difference between these two polyphenols is chemical structure and thus appropriate analytical tools were applied for each subgroup. For the condensed tannins, water-extracted, solvent-extracted, bisulphited water-extracted quebracho tannins and water-extracted and bisulphited water-extracted mimosa tannins were considered. In this case for method development the cacao tannin was used as a reference. The hydrolysable tannins considered were all solvent extracts, from turkey gall, chestnut and tara tannin.

The results obtained in this thesis can be divided into two sections, each dedicated to the type of tannin that was considered, a summary is offered:

1) Analysis of the condensed tannins carried out by averaging spectroscopic techniques revealed the difference in chemical composition and molar mass information. The extracts obtained from the quebracho and mimosa woods were analysed by ^{13}C NMR and compared with the cacao tannin and each other. A relative comparison of the predominant phenolic ring structures was possible; the mimosa tannin is composed of the highest proportion of prodelphinidin (PD) units compared to the other two tannins. The results also showed that mimosa is more prone to form C4-C6 bonds whereas the quebracho tannin tends to form C4-C8 bonds. By looking at the relative intensity of the peaks assigned to the interflavonoid bond a qualitative comparison of the molar mass was possible. The quebracho tannin (bisulphited water-extracted) indicated the highest molar mass since it showed the highest intensity of the signal given by C4-C8 interflavonoid bond. The peaks observed at 178 ppm in the ^{13}C NMR spectrum indicated that some degradation was present and could have been as a result of the extraction method. The peaks in the same region in the cacao tannin were attributed to the presence of galloyl residues that are present in the extract. MALDI-TOF MS confirmed the structural assignments made by ^{13}C NMR. Additional information was obtained with regards to the molar masses of the tannins. MALDI-TOF also revealed directly the differences in structure between the tannins analysed. Quebracho tannins obtained by different extraction methods were compared and it was quite clear that the extraction method alters the chemical composition and molar mass of the extract due to hydrolysis that occurs in the process. Although a vast amount of information was obtained in MALDI-TOF, one of its drawbacks is that it cannot distinguish between two isomers with the same mass but different chemical composition, and as a result, ambiguous structural assignments were made. In order to determine the actual structures present, MALDI-CID was employed. This technique enabled monomer sequence determination in the oligomer and this enabled positive assignments of the isobaric structures observed in MALDI-TOF. The CID experiments revealed two important facts, (1) although MALDI-TOF is a soft ionisation technique, tannin molecules undergo some fragmentation during analysis (2) the mode of fragmentation observed in CID mode is dependent on chemical structure. The experiments performed in CID thus revealed the significance of the laser intensity used in the analysis of tannins. The ^{13}C NMR analysis only provided average information on molar mass and chemical composition. In order to obtain more specific information liquid chromatographic separations were conducted. A HILIC method developed for the analysis of cacao tannin was applied to separate extracts from quebracho and mimosa wood. The gradient applied gave good separation of the molecules at lower elution volumes; however, a

'hump' was observed at higher elution volumes. The separation did not show good resolution unlike in the case of the cacao tannin, as expected. This presented a challenge in determining the elution order of the peaks. In order to tackle this problem, the analytical separation was up-scaled to a preparative fractionation under the same elution conditions. The fractions that were collected were then analysed by MALDI-TOF MS which revealed that the separation was indeed by size; however, a chemical composition separation was superimposed on it. This hyphenation also indicated the difference in chemical composition of the various extracts. The peaks that were presented in the MALDI-TOF spectrum revealed mainly a chemical composition separation. And mostly lower molar mass oligomers up to tetramers were detected. The galloyted molecules observed in the solvent-extracted quebracho and cacao tannin seem to be retained longer and elute with the higher molar mass fraction. A peculiarity was observed in this analysis though. The MALDI-TOF seems to favour ionisation of some molecules over others; the reason for this conclusion is that some of the peaks observed in the bulk sample MALDI-TOF spectrum were not observed in the fraction, in some cases the intensity was really low compared to what was expected. This loss of information was attributed to the fact that some oligomer information is lost during re-dissolving in the fractionation step and the fact that MALDI-TOF could cause fragmentation of larger molecules. Since the HILIC separation did not separate the molecules solely by size, SEC was utilised for this purpose. Non-derivatised tannin samples were analysed and their relative molar masses relative to PMMA were determined. The analysis revealed the quebracho tannin has the highest molar mass. The degree of polymerisation is comparable and varies from 5-20 units. The maleic anhydride modified quebracho tannin showed the most complex structure and this is a result of the modification.

2) For the hydrolysable tannins, the structure and molar masses of the commercial tara tannin, turkey gall tannin and chestnut tannin were determined by MALDI-TOF MS. The tara and turkey gall tannins were confirmed to be gallotannins. Although this is the case the basic building blocks of the tannins were different. Analysis of the tara tannin revealed that it contained some quinone structures and these were not present in the case of the turkey gall tannin. The chestnut tannin was revealed to be an ellagitannins that consists of complex structures in the extract. The mode of extraction was shown to vary the content of the chestnut tannin to a great extent. Solvent extraction seems to favour the oligomeric carbohydrate chains and the

oxidation products, vescalin/castalin and vescalagin/castalagin were not detected. MALDI CID experiments were conducted on various selected oligomers of the tannin extracts. The CID experiment revealed the monomer sequence in this case and confirmed the proposed structure of the tara tannin. The oligomer structure whereby the gallic acid residues are attached linearly, to a terminating molecule, was found to fragment in a simple manner, as was the case in the tara tannin. The turkey gall tannin which consists of oligomers of glucose attached to gallic acid residues. The presence of the sugar molecule quite clearly altered the fragmentation pattern and it was shown that the gallic acid residues directly attached to glucose were not likely to fragment. The exact oligomer structure of an oligomer present in the chestnut tannin was determined. Molar mass distribution information was obtained by SEC and this method proved its usefulness in being able to separate all the tannin extracts regardless of the chemical composition. A NP-LC method was developed for analysis of gallotannins, and by inference this method was able to indicate the differences in molar mass and chemical composition of the two extracts. The tara tannin seems to be composed of a higher proportion of polymeric molecules as compared to the turkey gall tannin.

In future, more condensed tannins require analysis by the MALDI CID technique in order to remove the ambiguity inherent in the structural assignment of MALDI-TOF peaks. The supposedly galloyted structures in the cacao tannin can be compared to the hydrolysed fragments observed in the modified quebracho tannin. These two chemical differences give the same repeat unit of 152 Da in MALDI-TOF. The same experiment needs to be carried out for other hydrolysable tannins in order for a comparative study to be possible. The SEC method developed can be improved by using the fractions from the HILIC experiment and performing a calibration instead of using PMMA standards. The SEC method for analysis of hydrolysable tannins may require a different mobile phase such that fractionation and subsequent MALDI-TOF analysis is made possible. The NP-LC method developed for the hydrolysable tannins can be further coupled to MALDI-TOF and online ESI-MS such that the elution order, chemical composition and molar mass of the tannin extracts may be determined. A liquid chromatographic method that is able to analyse both the ellagitannins and gallotannins needs to be developed.

In conclusion, it was revealed that several spectroscopic and liquid chromatography methods are required in order to fully characterise oligomeric tannins. The mode of extraction plays an important role in determining the structure, chemical and molar mass composition of each extract. The analytical methods that are used are indeed structure specific and care needs to be taken in making use of the correct method. The most robust way to analyse the tannins is by combining of liquid chromatographic methods with MS techniques. In order to combine the information obtained in each dimension, dedicated experimental parameters are necessary.

References

[1] H. Pasch, A. Pizzi, K. Rode, Polymer 42 (2001) 7531.

[2] A. Pizzi, H. Pasch, K. Rode, S. Giovando, J Appl Polym Sci 113 (2009) 3847.